


```
1090      1100      1110      1120      1130      1140      1150
      X          10          20          X
      KLGFFRSARRRRPGLDTPKYLE
      |||||
      LLLALLVIALRKLGFFRSARRRRPGLDTPKYLE
      1160      X 1170      1180      X
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 12:52:17 ; Search time 1.2067 Seconds
(without alignments)
1913.143 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1188
Perfect score: 125
Sequence: 1 KLGFRRSARRRRPGLDTPPKVLE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78:*

1: Pir1: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	52	41.6	956	JS0685	glutamate receptor
2	49	39.2	447	A84258	trans lesion repair
3	48	38.4	373	E70338	probable aspartate
4	48	38.4	528	T41362	hypothetical prote
5	48	38.4	798	T38490	hypothetical prote
6	48	38.4	3588	T40485	surfactin syntheta
7	47	37.6	231	C83525	TolQ protein PA096
8	46	36.8	120	A10961	conserved hypothet
9	46	36.8	519	S78089	G-protein signalin
10	46	36.8	536	T67945	3',5'-cyclic-nucle
11	46	36.8	589	S42047	aspartate-tRNA lig
12	46	36.8	708	H95367	probable histidine
13	46	36.8	712	S71626	3',5'-cyclic-nucle
14	46	36.8	1370	G84384	DNA polymerase typ
15	45	36.0	267	B33904	cAMP phosphodieste
16	45	36.0	485	PC4427	Nck, Ash and phosph
17	45	36.0	861	S77403	hypothetical prote
18	45	36.0	1435	D96693	protein Putative A
19	45	36.0	1607	G87259	conserved hypothet
20	44.5	35.6	244	S44822	F44E2.3 protein -
21	44.5	35.6	499	A12449	hypothetical prote
22	44	35.2	283	E95416	hypothetical prote
23	44	35.2	289	C87019	hypothetical prote
24	44	35.2	302	JH0572	hypothetical prote
25	44	35.2	344	A48990	transcriptional regu
26	44	35.2	344	T25567	hypothetical prote
27	44	35.2	385	A24247	hypothetical prote
28	44	35.2	406	G70639	hypothetical prote
29	44	35.2	760	JQ1383	hypothetical 84k p

30 44 35.2 1214 2 S28499 probable finger pr
31 44 35.2 1234 2 A34911 band 3-related pro
32 44 35.2 1237 2 A31789 band 3-related pro
33 44 35.2 1237 2 A56764 band 3-related pro
34 44 35.2 1450 2 A84780 probable ABC trans
35 44 35.2 1810 2 T31092 probable voltage-g
36 44 35.2 2649 2 T51023 hypothethical prote
37 43 34.4 340 2 C95940 probable allantoic
38 43 34.4 348 2 G83976 cytochrome caa3 ox
39 43 34.4 353 2 F75581 hypothethical prote
40 43 34.4 417 2 F95953 probable glycosylt
41 43 34.4 437 2 A72498 probable DNA/panto
42 43 34.4 453 2 G69070 coenzyme F390 synt
43 43 34.4 462 2 G87500 hypothethical prote
44 43 34.4 584 2 B53109 3',5'-cyclic-nucle
45 43 34.4 672 2 I61259 3',5'-cyclic-nucle

ALIGNMENTS

RESULT 1

JS0685
glutamate receptor chain KA-1 precursor - rat
N:Alternate names: kainate receptor kal
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 20-Aug-1999
C:Accession: JS0685; S18266; S19154
R:Herb, A.; Burnashev, N.; Werner, P.; Sakmann, B.; Wisden, W.; Seeburg, P.H.
Neuron 8, 775-785, 1992
A:Title: The KA-2 subunit of excitatory amino acid receptors shows widespread expression
A:Reference number: JH0592; MUID:92232306; PMID:11373632
A:Accession: JS0685
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-956 <HER>
A:Cross-references: ENBL:X59996; NID:957614; PIDN:CAA42615.1; PID:G57615
R:Werner, P.; Voigt, M.; Keinonen, K.; Wisden, W.; Seeburg, P.H.
Nature 351, 742-744, 1991
A:Title: Cloning of a putative high-affinity kainate receptor expressed predominantly in
A:Reference number: S18266; MUID:91287798; PMID:1648176
A:Accession: S18266
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-915, WHAHPVGVPAALPGAGSTVASAQRGEVGVGDHQQQRA' <WER>
A:Cross-references: ENBL:X59996
C:Superfamily: glutamate receptor; glutamate receptor homology
C:Keywords: neurotransmitter receptor; transmembrane protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-556/Product: glutamate receptor chain KA-1 #status predicted <MAT>
F:420-538/Domain: glutamate receptor homology <GRH>
F:547-566/Domain: transmembrane #status predicted <TM1>
F:591-609/Domain: transmembrane #status predicted <TM2>
F:620-638/Domain: transmembrane #status predicted <TM3>
F:805-825/Domain: transmembrane #status predicted <TM4>

Query Match 41.6%; Score 52; DB 2; Length 956;

Best Local Similarity 68.8%; Pred. No. 11;

Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 RRRRPPGLDTPPKVLE 24

||||| ||||| |||||

DB 862 RRRRSGGLPPQPPVLE 877

RESULT 2

A84258
trans lesion repair [imported] - Halobacterium sp. NRC-1

C:Species: Halobacterium sp. NRC-1

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: A84258

R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky,
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl

Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of *Halobacterium* species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: A84258
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-447 <STO>
A:Cross-references: GB:AE004437; NID:g10580569; PIDN:AAG19429.1; GSPDB:GN00138
C:Genetics:
A:Gene: yqjH

Query Match	39.2%	Score 49;	DB 2;
Best Local Similarity	50.0%	Pred. No. 14;	
Matches 10;	Conservative	2;	Mismatches 8;
			Indels 0;
			Gaps 0;

Qy	4	FFRSARRRREPGLDPTPKVL	23
		: :	
Db	62	FVASCERREPALRGAPVVV	81

RESULT 3
E70338
Probable aspartate transaminase [EC 2.6.1.1] aspC3 [similarity] - Aquifex aeolicus
C:Species: Aquifex aeolicus
C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 17-Mar-2000
F:Accession: E70338
F:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.

Nature 392, 353-358, 1998
A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: E70338
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-373 <AOF>

A;Cross-references: GB:AE000689; NID:g2983082; PIDN:AAC06689.1; PID:g2983083; GB:AE000695
A;Experimental source: strain VF5

C:Genetics:
A:Gene: aspc3
C:Superfamily: aspartate transaminase
C:Keywords: aminotransferase; phosphoprotein; pyridoxal phosphate
F.222/Binding site: pyridoxal phosphate (lys) (covalent) #status predicted

Query Match 38.4%; Score 48; DB 2; Length 373;
Best Local Similarity 66.7%; Pred. No. 17;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY	13	EPGLDPTPKVLE	24
		: : : :	
Dd	32	EPDLEPSPKVMK	43

RESULT 4
T41362
hypothetical protein SPC4G3.14 - fission yeast (*Schizosaccharomyces pombe*)
C:Species: *Schizosaccharomyces pombe*
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 02-Sep-2000
C:Accession: T41362
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Hiltbert, H.; Duesterhoeft, A.
submitted to the EMBL Data Library, March 1998

A:Reference number: Z21918
A:Accession: T41362
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-528 (WOO)

A: Cross-references: EMBL:Z37052; PIDN: CAB09769.1; GSPDB:GN00068; SPDB:SPCC4G3.14
A: Experimental source: strain 972b-; cosmid c4G3

Genetics: [Genetics](#)
 Gene: [SPDB:SPCC4G3.14](#)
 Map position: [3](#)
 Superfamily: [heat shock protein dnaJ](#); [dnaJ amino-terminal homology](#)

F;86-150/Domain: dnaJ amino-terminal homology <DNJ>

```
Query Match      38.4%; Score 48; DB 2; Length 528;
Best Local Similarity 52.6%; Pred. NO. 24;
Matches 10; Conservative 1; Mismatches 8; Indels
```

QY 3 GFFSARRRRRPGLDTPK 21
|||: ||| | |||
Db 504 GFFKRAFRRLHPDEDQNEK 522

RESULT 5

hypothetical protein SPAC29H12.02c - fission yeast (Schizosaccharomyces pombe)
C-Species: Schizosaccharomyces pombe
C-Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C-Accession: T38490
R-Genies, S.; Churcher, C.M.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997

A.Reference number: Z21797
 A.Accession: T88490
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: DNA
 A.Residues: 1-798 <GEN>
 A.Cross-references: EMBL:Z99164; PIDN:CAB16247.1; GS
 A.Experimental source: strain 972h; cosmid c29B12
 C.Genetics:
 A.Gene: SPDB:SPAC29B12.02c
 A.Map position: 1

Query Match 38.4%; Score 48; DB 2; Length 798;
Best Local Similarity 50.0%; Pred. No. 37;

	Matches	8; Conservative
QY	9	RRRREPGLDPTPKVLE 24
		: : : : : : :
Db	355	RORRKGIDETSKIIE 370

RESULT 6

I40485
I40485
surfactant synthetase component I - *Bacillus subtilis*
N:Alternate names: competence protein srfA; surfactin production protein srfA; surfact
C:Species: *Bacillus subtilis*
C:Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 03-Nov-2000
C:Release: I40485; B69718; S35117; A37323; S45967; A43705; S34985
C:Accession: I40485; B69718; S35117; A37323; S45967; A43705; S34985
C:Author: R., Cosmina, P.; Rodriguez, F.; de Ferra, F.; Grandi, G.; Prego, M.; Venema, G.; van S
C:J: Mol. Microbiol. 8, 821-831, 1993
A:Title: Sequence and analysis of the genetic locus responsible for surfactin synthesis
A:Reference number: I40485; MUID:93360813; PMID:8355609
A:Accession: I40485
A:Status: translated from GB/EMBL/DBJ

A: Molecule type: DNA
A: Residues: 1-3588 <RES>
A: Cross-references: EMBL:X70356; NID:g396480; PIDN:CAA49816.1; PID:g396481
A: Experimental source: strain W168 derivative of JH642

Nature 390, 249-256, 1997

Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler, R.; Gogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, R.; Rigier, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, J.; Schreuter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Szor, A.; Tamakoshi, A.; Tanaka, T.; Terstra, P.; Tognoni, A.; Tosca, V.; Uchiyama, K.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshioka, H.F.; Zmstein, E.; Yoshiokawa, H.; Danchin, A.

Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.

Reference number: A69580; PMID:98044033; PMID:9384377

Accession: B69718

A>Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-3588 <UN>

A:A:Cross-references: GB:Z90105; GB:AL009126; NID:g2632457; PIDN:CAB12142.1; PID:g2632634

A:Experimental source: strain 168

A:Experimental source: Fujishima, Y.; Corbell, N.; D'Souza, C.; Nakano, M.M.; Zuber, P.; Yamane, K.

A:Nucleic acid sequence: 21, 93-97, 1993

A:Title: Nucleotide sequence of 5' portion of srfA that contains the region required for

A:Reference number: S35517; MUID:93181186; PMID:8441623

A:Accession: S35517

A:A>Status: nucleic acid sequence not shown; significant sequence differences

A:Molecule type: DNA

A:A:Cross-references: EMBL:D13262; NID:g216345; PID:g216346

A:Experimental source: strain 168 trpC2

A:Note: protein sequence not complete, the nucleotide sequence was submitted to the EMBI

A:R.Nakano, M.M.; Magnuson, R.; Myers, A.; Curry, J.; Grossman, A.D.; Zuber, P.

A:J. Bacteriol. 173, 1770-1778, 1991

A:Title: srfA is an operon required for surfactin production, competence development, an

A:Reference number: A37323; MUID:91154134; PMID:1847909

A:Accession: A37323

A:A>Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Molecule type: DNA

A:A:Residues: 1-46, 'EV', 49-145, 'I', 147-150, 'L', 152, 'AN', 155-280, 'T', 282-307, 'SF', 310-384 <

A:Cross-references: GB:M59339; NID:g143366

A:R.Nakano, M.M.; Xia, L.; Zuber, P.

A:J. Bacteriol. 173, 5487-5493, 1991

A:Title: Transcription initiation region of the srfA operon, which is controlled by the

A:Reference number: A43705; MUID:91358326; PMID:1715856

A:Contents: annotation

A:R.Fabret, C.; Quentin, Y.; Guiseppi, A.; Busuttill, J.; Haiech, J.; Denizot, F.

A:Reference number: S46967

A:Accession: S46967

A:Molecule type: DNA

A:Residues: 3249-3271, 'A', 3273-3316, 'R', 3318-3451, 'Y', 3453-3483, 'DE', 3486-3487, 'DAGL', 34

A:Cross-references: EMBL:X72672; NID:g516359; PIDN:CAAS1222.1; PID:g516359

A:Experimental source: strain 168 trpC2

A:Comment: This enzyme is one of several in the multienzyme complex that synthesizes the

A:Genetics:

A:Gene: srfA; srfA1

A:Note: srfA is the first gene of the srfA operon

A:Superfamily: surfactin synthetase; acetate-CoA ligase homology; acyl carrier protein h

A:Keywords: antibiotic biosynthesis; carrier protein; phosphopantetheine; phosphoprotein

A:F507-956/Domain: acetate-CoA ligase homology <ACL1>

A:F974-1042/Domain: acyl carrier protein homology <ACP1>

A:F1043-1488/Domain: repeat <RPT1>

A:F1549-1993/Domain: acetate-CoA ligase homology <ACL2>

A:F2011-2079/Domain: acyl carrier protein homology <ACP2>

A:F2080-2527/Domain: repeat <RPT2>

A:F2589-3025/Domain: acetate-CoA ligase homology <ACL3>

A:F3042-3109/Domain: acyl carrier protein homology <ACP3>

A:F1006,2043,3074/Binding site: phosphopantetheine (Ser) (covalent) #status predicted

Query Match 38.4%; Score 48; DB 2; Length 3588;

Best Local Similarity 45.5%; Pred. No. 1.7e+02;

Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 3 GFFRSARRRRPGDPTPKVLE 24

DB 1093 GFFYSKAGARISGINTPRVIE 1114

RESULT 7

C83525

TolQ protein PA0969 [imported] - Pseudomonas aeruginosa (strain PAO1)

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Nov-2001

C:Accession: C83525

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lam,

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A>Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: CB3525
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-231 <STO>
A:Cross-references: GB:AEOC4530; GB:AEOC4091; NID:g9946865; PIDN:AAG04358.1; GSPDB:GN001
A:Experimental source: strain PA01

C:Genetics:
A:Gene: tolQ; PA0969
C:Superfamily: biopolymer transport protein

Query Match 37.6%; Score 47; DB 2; Length 231;
Best Local Similarity 44.4%; Pred. No. 15;
Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRRREPGLDP 18
 : ||| : ||| : ||| : |||
Db 89 RAGFKFESRLRQQGVDP 106

RESULT 8

AI0961 conserved hypothetical protein STY3975 [imported] - Salmonella enterica subsp. enterica serovar Typhimurium
C:Species: Salmonella enterica subsp. enterica serovar Typhi
A>Note: This species has also been called Salmonella typhi
C>Date: 09-Nov-2001 #sequence_revision 09-Nov-2001.#text_change 18-Nov-2002
C:Accession: AI0961
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; White, N.; Farrar, L.; Davies, R.M.; Dowd, L.; White, N.; Farrar, L.; S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parkhill, J.; Quail, M.A.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; White, N.; Farrar, L.; Davies, R.M.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, C.; White, N.; Farrar, L.; S.; Moule, S.; O'Garra, P.
A>Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium
A:Reference number: AB0502; MUID:21534947; PMID:11677608
A:Accession: AI0961
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD03189.1; PID:g16504821; GSPDB:GN00176
C:Genetics:
A:Gene: STY3975

Query Match 36.8%; Score 46; DB 2; Length 120;
Best Local Similarity 61.5%; Pred. No. 11;
Matches 8; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 6 RSARRRRREPGLDP 18
 ||| | : ||| : |||
Db 5 RXARRNDPGLQP 17

RESULT 9

G-protein signaling regulator RGS3 - human
C:Species: Homo sapiens (man)
C>Date: 04-Dec-1997 #sequence_revision 12-Dec-1997 #text_change 24-Nov-1999
C:Accession: S78089; S68436
R:Druey, K.
submitted to the EMBL Data Library, May 1995
A:Reference number: S78089
A:Accession: S78089
A:Molecule type: mRNA
A:Residues: 1-519 <DRU>
R:Druey, K.M.; Blumberg, P.M.; Kang, V.H.; Kehrl, J.H.
Nature 379, 742-746, 1996
A>Title: Inhibition of G-protein-mediated MAP kinase activation by a new mammalian gene product
A:Reference number: A58012; MUID:96178495; PMID:8602223
A:Accession: S68436
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-355, 'K', 357-519 <DRW>

A:Cross-references: EMBL:U27655
 A:Experimental source: tonsil
 A>Note: the sequence from Fig. 1b is inconsistent with that from Fig. 1a in having 356-1

Query Match 36.8%; Score 46; DB 2; Length 519;
 Best Local Similarity 52.4%; Pred. No. 47;
 Matches 11; Conservative 0; Mismatches 8; Indels 2; Gaps 1;

QY 1 KLGFFSARRRRPGLDPTPK 21
 |||||
 DB 355 KLGIFR--RRNESPCAPPACK 373

RESULT 10
 167945
 3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17) - rat (fragment)
 N:Alternate names: CAMP phosphodiesterase 1
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 03-Nov-2000
 C:Accession: I67945; A33904
 R:Bolger, G.B.; Rodgers, L.K.; Riggs, M.
 Gene 149, 237-244, 1994
 A:Title: Differential CNS expression of alternative mRNA isoforms of the mammalian genes
 A:Reference number: I53865; MUID:95047482; PMID:7958996
 A:Accession: I67945
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-536 <RES>
 A:Cross-references: GB:I27061; NID:9436011; PIDN:AAA56858.1; PID:9436012
 R:Swinnen, J.V.; Joseph, D.R.; Conti, M.
 Proc. Natl. Acad. Sci. U.S.A. 86, 5325-5329, 1989
 A:Title: Molecular cloning of rat homologues of the Drosophila melanogaster dunce cAMP P
 A:Reference number: A33904; MUID:89315790; PMID:2546153
 A:Accession: A33904
 A>Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 230-496 <SWI>
 A:Cross-references: GB:M25347; GB:M28410
 C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase 1B, calmodulin-dependent; 3',5'
 C:Keywords: alternative splicing; cAMP binding; phosphoric diester hydrolase
 F:253-481/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNPD>

Query Match 36.8%; Score 46; DB 2; Length 536;
 Best Local Similarity 52.9%; Pred. No. 49;
 Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 FFRSARRRRPGLDPTPK 20
 |||||
 DB 434 FFQGGRRERESGLDISP 450

RESULT 11
 S42047
 aspartate-tRNA ligase (EC 6.1.1.12) - Mycobacterium leprae
 N:Alternate names: aspartyl-tRNA synthetase
 C:Species: Mycobacterium leprae
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: S42047
 R:Spierings, E.H.T.; Wiele, B.; Thole, J.E.R.
 Submitted to the EMBL Data Library, February 1994
 A:Reference number: S42047
 A:Accession: S42047
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-589 <SPI>
 A:Cross-references: EMBL:X77655; NID:9454973; PIDN:CAA54735.1; PID:9581332
 C:Genetics:
 A:Gene: ASPS
 A:Start codon: GTG
 C:Function:
 A:Description: activates amino acid and transfers it to specific tRNA molecule
 A:Pathway: protein biosynthesis
 C:Superfamily: lysine-tRNA ligase

C:Keywords: aminoacyl-tRNA synthetase; ATP; ligase; protein biosynthesis
 Query Match 36.8%; Score 46; DB 1; Length 589;
 Best Local Similarity 44.4%; Pred. No. 54;
 Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

QY 7 SARRRRPGLDPTPKVLE 24
 |||||
 DB 572 TAQQRKESGIDTKPEVE 589

RESULT 12
 H95367
 probable histidine kinase, chemotaxis (EC 2.7.3.-) [imported] - Sinorhizobium meliloti (C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 22-Jun-2003
 C:Accession: H95367
 R:Barrett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
 ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: H95367
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-708 <KUR>
 A:Cross-references: GB:AE006469; PIDN:AAK65506.1; PID:gl4523979; GSPDB:GN00165
 A:Experimental source: strain 1021, megaplasmid pSymA
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 P.; D. Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: SMal558
 A:Genome: plasmid
 C:Superfamily: chemotaxis hybrid-type signal transduction histidine kinase, PrZe type; r
 C:Keywords: phosphotransferase

Query Match 36.8%; Score 46; DB 2; Length 708;
 Best Local Similarity 60.0%; Pred. No. 65;
 Matches 9; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 8 ARRRRRPGLDPTPKV 22
 |||||
 DB 204 ARKMRAPGSDPTGV 218

RESULT 13
 S71626
 3',5'-cyclic-nucleotide phosphodiesterase (EC 3.1.4.17), cAMP-specific - human
 N:Alternate names: 3',5'-cyclic AMP phosphodiesterase
 C:Species: Homo sapiens (man)
 C>Date: 27-Nov-1997 #sequence_revision 12-Dec-1997 #text_change 17-Nov-2000
 C:Accession: S71626; I61356
 R:Engels, P.; Sullivan, M.; Mueller, T.; Luebbert, H.
 FEBS Lett. 358, 305-310, 1995
 A:Title: Molecular cloning and functional expression in yeast of a human cAMP-specific p
 A:Reference number: S71626; MUID:95145731; PMID:7843419
 A:Accession: S71626
 A:Molecule type: DNA
 A:Residues: 1-712 <ENG>
 A:Cross-references: EMBL:Z46632; NID:g727222; PIDN:CAA86601.1; PID:g727223
 A:Experimental source: substantia nigra
 R:Bolger, G.; Michaeli, T.; Martins, T.; St. John, T.; Steiner, B.; Rodgers, L.; Riggs,
 Mol. Cell. Biol. 13, 6588-6571, 1993
 A:Title: A family of human phosphodiesterases homologous to the dunce learning and memor
 A:Reference number: A54442; MUID:94019330; PMID:8413254
 A:Accession: I61356

A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 462-712 <RES>
A:Cross-references: GB:L20968; NID:g347125; PIDN:AAA03591.1; PID:g347126
C:Genetics:
A:Gene: HSPD4C1
C:Function:
A>Description: CAMP hydrolysis; converts CAMP to non-cyclic AMP
A:Pathway: cyclic nucleotide metabolism
A>Note: expressed in various tissues but not in cells of the immune system
C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase 1B, calmodulin-dependent; 3',5'
C:Keywords: phosphoric diester hydrolase
F:387-615/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match 36.8%; Score 46; DB 2; Length 712;
Best Local Similarity 52.9%; Pred. No. 65;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 4 FFRSARRRRPGLDPTP 20
||: ||| ||| ||| :
Db 568 FFOQGDREERESGLDISP 584

RESULT 14
G84384
DNA polymerase type II large chain [imported] - Halobacterium sp. NRC-1
C:Species: Halobacterium sp. NRC-1
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
C:Accession: G84384
R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.
; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jabl
Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of Halobacterium species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: G84384
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1370 <STO>
A:Cross-references: GB:AE004437; NID:g10581749; PIDN:AAG20443.1; GSPDB:GN00138
C:Genetics:
A:Gene: polA2

Query Match 36.8%; Score 46; DB 2; Length 1370;
Best Local Similarity 55.6%; Pred. No. 1.3e+02;
Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 FRSARRRRPGLDPTPKV 22
| | | | | | | | | |
Db 20 FEVAARERGEDPTFNV 37

RESULT 15
B33904
CAMP phosphodiesterase 2 - rat (fragment)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 02-Sep-2000
C:Accession: B33904
R;Swinnen, J.V.; Joseph, D.R.; Conti, M.
Proc. Natl. Acad. Sci. U.S.A. 86, 5325-5329, 1989
A:Title: Molecular cloning of rat homologues of the Drosophila melanogaster dunce cAMP P
A:Reference number: A33904; MUID:89315790; PMID:2546153
A:Accession: B33904
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-267 <SWI>
A:Cross-references: GB:M25348
C:Superfamily: 3',5'-cyclic-nucleotide phosphodiesterase homology
F:24-252/Domain: 3',5'-cyclic-nucleotide phosphodiesterase homology <CNP>

Query Match 36.0%; Score 45; DB 2; Length 267;
Best Local Similarity 47.1%; Pred. No. 34;

Matches 8; Conservative 4; Mismatches 5; Indels 0; Gaps 0;
QY 4 FFRSARRRRPGLDPTP 20
||: ||| ||| ||| :
Db 205 FFOQAHRRERGMWISP 221

Search completed: September 21, 2004, 13:05:04
Job time : 5.2067 secs

This page blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 12:44:02 ; Search time 0.703911 Seconds
(without alignments)
1775.344 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1189
Perfect score: 125
Sequence: 1 KLGFRSARRRRPGLDTPKYLE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues
Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	1189	1 ITAH_HUMAN	Q9ukx5 homo sapien
2	52	41.6	956	1 GLK4 RAT	Q01812 rattus norv
3	49	39.2	411	1 DPO4 HALN1	Q9hqt4 halobacteri
4	48	38.4	358	1 SRFI_BACSU	P27206 bacillus su
5	47	37.6	231	1 TOLG_PSEAE	P50598 pseudomonas
6	47	37.6	262	1 ORC6_MOUSE	Q9wuj8 mus musculu
7	46	36.8	519	1 RGS3_HUMAN	P49796 homo sapien
8	46	36.8	536	1 CN4C RAT	P14644 rattus norv
9	46	36.8	589	1 SYD_MYCLE	P36429 mycobacteri
10	46	36.8	712	1 CN4C_HUMAN	Q08493 homo sapien
11	46	36.8	1370	1 DP2L_HALN1	Q9hmx8 halobacteri
12	45	36.0	485	1 SOC6_MOUSE	Q8vht2 mus musculu
13	45	36.0	579	1 SYFB_THETN	Q8r9c7 thermonaer
14	45	36.0	794	1 SYFB_THETN	P34433 caenorhabdi
15	44.5	35.6	244	1 YL53_CABEL	O69583 mycobacteri
16	44	35.2	289	1 QRC8_MYCLE	Q91127 streptomyce
17	44	35.2	328	1 MPR2_STRCO	P43161 streptomyce
18	44	35.2	344	1 MPR2_STRCO	Q92866 rickettsia
19	44	35.2	554	1 Y478_RICPR	Q8twx0 methanopyru
20	44	35.2	564	1 CABI_METKA	Q14137 homo sapien
21	44	35.2	746	1 BOPI_HUMAN	Q8p726 xanthomonas
22	44	35.2	791	1 SYFB_XANCP	O8pj65 xanthomonas
23	44	35.2	792	1 SYFB_XANCP	O8bjm5 mus musculu
24	44	35.2	825	1 NLG3_MOUSE	Q9nt94 homo sapien
25	44	35.2	848	1 NLG3_HUMAN	Q62889 rattus norv
26	44	35.2	848	1 NLG3_RAT	Q63679 rattus norv
27	44	35.2	1214	1 TSGA_RAT	P23347 rattus norv
28	44	35.2	1234	1 B3A2_RAT	P13808 mus musculu
29	44	35.2	1237	1 B3A2_MOUSE	P48746 oryctolagus
30	44	35.2	1237	1 B3A2_RABIT	Q9z058 cavia porce
31	44	35.2	1238	1 B3A2_CAVPO	Q62261 mus musculu
32	44	35.2	2363	1 SPFO_MOUSE	P47810 mus musculu
33	43.5	34.8	646	1 WEEL_MOUSE	

RESULT 1

ID	ITAH_HUMAN	STANDARD;	PRT;	1189 AA.
AC	Q9UKX5; Q9UKQ1;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Integrin alpha-11 precursor.			
GN	ITGA11.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal heart, and Osteoblast;			
RX	MEDLINE=99417679; PubMed=10486209;			
RA	Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,			
RA	Wang S.-X., Morris C.M., Kriessens G.W.;			
RT	"Cloning, sequence analysis, and chromosomal localization of the novel			
RT	human integrin alpha11 subunit (ITGA11).";			
RL	Genomics 60:179-187(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal muscle, and Uterus;			
RX	MEDLINE=99395147; PubMed=10464311;			
RA	Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;			
RT	"cDNA Cloning and Chromosomal Localization of Human alpha(11)			
RT	Integrin. A collagen-binding, i domain-containing, beta(1)-associated			
RT	integrin alpha-chain present in muscle tissues.";			
RL	J. Biol. Chem. 274:25735-25742(1999).			
RN	[3]			
RP	SEQUENCE OF 954-1188 FROM N.A.			
RC	TISSUE=Fibroblast;			
RA	Andreu N., Estivill X., Escarceller M., Sumoy L.;			
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.			
CC	-!- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.			
CC	-!- SUBUNIT: HETERO-DIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11			
CC	ASSOCIATES WITH BETA-1.			
CC	-!- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-!- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND			
CC	HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO			
CC	LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO			
CC	FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,			
CC	SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN			
CC	PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.			
CC	-!- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING			
CC	FETAL MUSCLE CELLS (IN VITRO).			
CC	-!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS			
CC	WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.			
CC	-!- SIMILARITY: Belongs to the integrin alpha chain family.			
CC	-!- SIMILARITY: Contains 7 WGAP repeats.			
CC	-!- SIMILARITY: Contains 7 FGAP repeats.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			

Q63802 rattus norv
Q60840 homo sapien
P14270 rattus norv
Q08499 homo sapien
O14302 schizosacch
P11533 mycobacteri
P04920 homo sapien
Q9J187 mus musculu
Q88480 rattus norv
Q96rv3 homo sapien
Q8prt4 xanthomonas
Q8rtj5 xanthomonas


```

CC EMBL; X59996; CA242615.1; -
DR EMBL; U08257; AAA17830.1; -
DR PIR; J50685; J50685.
DR HSSP; P19491; IGR2.
DR InterPro; IPR001828; ANF receptor.
DR InterPro; IPR001320; Ion glu receptor.
DR InterPro; IPR001622; K-channel pore.
DR InterPro; IPR001508; NMDA receptor.
DR InterPro; IPR001311; SBP/Glu_receptor.
DR Pfam; PF01094; ANF_receptor; 1.
DR Pfam; PF00060; lig_chan; 1.
DR PRINTS; PR00177; NMDARECEPTOR.
DR SMART; SM00079; PBPE; 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Signal; Transmembrane.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 956
FT DOMAIN 21 545 GLUTAMATE RECEPTOR, IONOTROPIC KAINATE 4.
FT TRANSMEM 546 566 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 624 644 POTENTIAL.
FT TRANSMEM 805 825 POTENTIAL.
FT DOMAIN 862 865 POLY-ARG.
FT CARBOHYD 158 158 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 220 220 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 323 323 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 408 408 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 415 415 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 473 479 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 539 539 S -> F (IN REF. 2).
FT SEQUENCE 956 AA; 107223 MW; 970C9D72C8D74700 CRC64;

Query Match 41.6%; Score 52; DB 1; Length 956;
Best Local Similarity 68.8%; Pred. No. 5.6;
Matches 11; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 9 RRRRFGGLDPTPKVLE 24
||||| ||||| |||||
Db 862 RRRRSGGLPQPPVLE 877

RESULT 3
ID DPO4_HALN1 STANDARD; PRT; 411 AA.
AC Q9HQF4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE DNA polymerase IV (EC 2.7.7.7) (Pol IV).
DB DB OR VNG1014G.
OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=64091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=3050483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
RA Shukla H.B., Lasky S.R., Balliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
RA Leithauser B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenberger T.A., Peck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RA "Genome sequence of Halobacterium species NRC-1."
RA Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -!- FUNCTION: Poorly processive, error-prone DNA polymerase involved
CC in untargeted mutagenesis. Copies undamaged DNA at stalled
CC replication forks, which arise in vivo from mismatched or
CC misaligned primer ends. These misaligned primers can be extended
CC by polIV. Exhibits no 3'-5' exonuclease (proofreading) activity.

```

```

CC May be involved in translational synthesis (By similarity).
CC -!- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
CC + {DNA}(N).
CC -!- COFACTOR: Binds 2 magnesium ions per subunit (By similarity).
CC -!- SUBUNIT: Monomer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the DNA polymerase type-Y family.
CC -!- SIMILARITY: Contains 1 umuc domain.
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE005036; AAG19429.1; ALT_INIT.
CC HAMAP; MF 01113; -; 1.
CC InterPro; IPR001126; UMUC_like.
CC Pfam; PF00817; IMS; 1.
CC PROSITE; PS00173; UMUC; 1.
CC TRANSFERASE; DNA-directed DNA polymerase; DNA replication; DNA repair;
CC DNA-binding; Mutator protein; Magnesium; Complete proteome.
CC DOMAIN 18 211 UMUC.
CC SITE 27 27 SUBSTRATE DISCRIMINATION (BY SIMILARITY).
CC ACT_SITE 131 131 BY SIMILARITY.
CC METAL 22 22 MAGNESIUM (BY SIMILARITY).
CC METAL 130 130 MAGNESIUM (BY SIMILARITY).
CC SEQUENCE 411 AA; 43492 MW; CA152C89EBEAF81 CRC64;

Query Match 39.2%; Score 49; DB 1; Length 411;
Best Local Similarity 50.0%; Pred. No. 6.5;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 4 FFRSARRRREGGLDPTPKVL 23
||||| ||||| |||||
Db 26 FYASCRRRREPALRGAPVW 45

RESULT 4
SRF1_BACSU STANDARD; PRT; 3588 AA.
ID SRF1_BACSU STANDARD; PRT; 3588 AA.
AC P27206;
DT 01-AUG-1992 (Rel. 23, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Surfactin synthetase subunit 1.
GN SRF4A OR SRF4I OR SRF4 CR BSU03480.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=93181186; PubMed=8441623;
RA Fuma S., Fujishima Y., Corbell N., D'Souza C., Nakano M.M.,
RA Zuber P., Yamane K.;
RA "Nucleotide sequence of 5' portion of srfA that contains the region
RT required for competence establishment in Bacillus subtilis."
RL Nucleic Acids Res. 21:93-97(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168 / JH642;
RX MEDLINE=93360813; PubMed=8355609;
RA Cosmina P., Rodriguez F., de Perra F., Grandi G., Perego M.,
RA Venema G., van Sinderen D.;
RA "Sequence and analysis of the genetic locus responsible for surfactin
RT synthesis in Bacillus subtilis."
RL Mol. Microbiol. 8:821-831(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;

```

RX MEDLINE=97124189; PubMed=8969502;
 RA Yanane K., Kumano M., Kurita K.;
 RT "The 25 degrees-36 degrees region of the Bacillus subtilis
 chromosome: determination of the sequence of a 146 kb segment and
 identification of 113 genes.";
 RL Microbiology 142:3047-3056(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 AZevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.D., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Enrian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
 RA Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G.,
 RA Giuseppe G., Guy B.J., Haga K., Haiteh J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel D., Kasahara Y., Kluwe M.F., Itaya M., Jones L.,
 RA Joris B., Karamata D., Kasahara Y., Kluwe M.F., Itaya M., Jones L.,
 RA Kobayashi Y., Koetter P., Koningstein G., Krogh S., Kumano M.,
 RA Kurita K., Lapius A., Lardinois S., Lauber J., Lazarevic V.,
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,
 RA Noone D., O'Reilly M., Ogawa K., Ogawa K., Oudega B., Park S.H.,
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
 RA Rieger M., Rivolta C., Roche E., Roche B., Rose M., Sadale Y.,
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Solido B.,
 RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,
 RA Tosato V., Uchiyama S., Vandenbol M., Vanier F., Vassarotti A.,
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,
 RA Winters P., Wisat A., Yamamoto H., Yamane K., Yasunoto K., Yata K.,
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 subtilis.";
 RL Nature 390:249-256(1997).
 RN [5]
 RP SEQUENCE OF 1-460 FROM N.A.
 RX MEDLINE=91154134; PubMed=1847909;
 RA Nakano M.M., Magnusson R., Myers A.M., Curry J., Grossman A.D.,
 RA Zuber P.;
 RT "srfa is an operon required for surfactin production, competence
 development, and efficient sporulation in Bacillus subtilis.";
 RL J. Bacteriol. 173:1770-1778(1991).
 RN [6]
 RP SEQUENCE OF 1-64 FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=95219080; PubMed=7704255;
 RA Fujishima Y., Yamane K.;
 RT "A 10 kb nucleotide sequence at the 5' flanking region (32 degrees)
 of srfa of the Bacillus subtilis chromosome.";
 RL Microbiology 141:277-279(1995).
 RN [7]
 RP SEQUENCE OF 1-38 FROM N.A.
 RX MEDLINE=91358326; PubMed=1715856;
 RA Nakano M.M., Xia L., Zuber P.;
 RT "Transcription initiation region of the srfa operon, which is
 controlled by the comp-cmA signal transduction system in Bacillus
 subtilis.";
 RL J. Bacteriol. 173:5487-5493(1991).
 CC -!- FUNCTION: THIS PROTEIN IS A MULTIFUNCTIONAL ENZYME ABLE TO
 CC ACTIVATE AND POLYMERIZE THE AMINO ACIDS LEU, GLU, ASP AND VAL.
 CC ACTIVATION SITES FOR THESE AA CONSIST OF INDIVIDUAL DOMAINS.
 CC -!- COFACTOR: Contains 3 covalently bound phosphopantetheines.
 CC -!- PATHWAY: Cyclic peptide antibiotic surfactin biosynthesis.
 CC -!- SIMILARITY: Belongs to the ATP-dependent AMP-binding enzyme
 CC family.
 CC -!- SIMILARITY: Contains 3 acyl carrier domains.

CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D13262; BAA02522.1; -
 CC EMBL; X70356; CAA49816.1; -
 CC EMBL; D50453; BAA09882.1; -
 CC EMBL; Z39105; CAB12142.1; -
 CC EMBL; M59939; AAA22815.1; -
 CC EMBL; D30762; BAA21034.1; -
 CC EMBL; M64702; AAA22816.1; -
 CC PIR; I40485; I40485.
 CC HSP; P14687; I40485.
 CC Subtilisin; BG10168; srfaA
 CC InterPro; IPR000873; AMP-bind.
 CC InterPro; IPR001242; Condensatn.
 CC InterPro; IPR006163; Pp_bind.
 CC InterPro; IPR006162; Ppantne S.
 CC Pfam; PF00501; AMP-binding; 3.
 CC Pfam; PF00668; Condensation; 4.
 CC Pfam; PF00550; pp-binding; 3.
 CC PRINTS; PR00154; AMPBINDING.
 CC PROSITE; PS00012; PHOSPHOPANTETHEINE; 2.
 CC PROSITE; PS00455; AMP BINDING; 3.
 CC PROSITE; PS00075; ACP DOMAIN; 3.
 CC KW Ligase; Antibiotic biosynthesis; Phosphopantetheine; Sporulation;
 CC Multi-functional enzyme; Repeat; Complete proteome.
 CC FT REPEAT ? 1047 DOMAIN 1 (GLU-ACTIVATING).
 CC FT REPEAT ? 2084 DOMAIN 2 (LEU-ACTIVATING).
 CC FT REPEAT ? 3115 DOMAIN 3 (D-LEU-ACTIVATING).
 CC FT DOMAIN 976 1043 ACYL CARRIER (ACP) 1.
 CC FT DOMAIN 2013 2080 ACYL CARRIER (ACP) 2.
 CC FT DOMAIN 3044 3110 ACYL CARRIER (ACP) 3.
 CC FT BINDING 1006 1006 PHOSPHOPANTETHEINE (POTENTIAL).
 CC FT BINDING 2043 2043 PHOSPHOPANTETHEINE (POTENTIAL).
 CC FT BINDING 3074 3074 PHOSPHOPANTETHEINE (POTENTIAL).
 CC FT CONFLICT 146 146 M -> I (IN REF. 1 AND 5).
 CC FT CONFLICT 151 151 V -> L (IN REF. 1 AND 5).
 CC FT CONFLICT 165 165 G -> A (IN REF. 1).
 CC FT CONFLICT 281 281 Q -> T (IN REF. 1 AND 5).
 CC FT CONFLICT 460 460 D -> T (IN REF. 5).
 CC FT CONFLICT 540 540 A -> P (IN REF. 1).
 CC FT CONFLICT 562 562 Y -> I (IN REF. 1).
 CC FT CONFLICT 639 640 FT -> GS (IN REF. 1).
 CC FT CONFLICT 644 644 P -> R (IN REF. 1).
 CC FT CONFLICT 647 649 FLIP -> LIA (IN REF. 1).
 CC FT CONFLICT 1026 1026 Q -> L (IN REF. 1).
 CC FT CONFLICT 1065 1114 HWSSEHSSASSRWRSSAIHSSNSSEGFYSKAGARISG
 CC INTPRVIE -> SLAQRTYIVSQPDAGVGNPFAAIL
 CC EGPLDQKQERAFQGLIRHESLR (IN REF. 1).
 CC VC -> DS (IN REF. 1).
 CC L -> R (IN REF. 1).
 CC L -> V (IN REF. 1).
 CC T -> S (IN REF. 1).
 CC T -> D (IN REF. 1).
 CC A -> R (IN REF. 1).
 CC PAVFQMD -> LRCLSKWT (IN REF. 1).
 CC P -> L (IN REF. 1).
 CC Q -> H (IN REF. 1).
 CC HR -> QQ (IN REF. 1).
 CC L -> V (IN REF. 1).
 CC C -> S (IN REF. 1).
 CC S -> R (IN REF. 1).
 CC AV -> RC (IN REF. 1).
 CC MISSING (IN REF. 1).
 CC S -> T (IN REF. 1).
 CC TA -> SP (IN REF. 1).
 CC N -> P (IN REF. 1).
 CC CONFLICT 1130 1131
 CC CONFLICT 1162 1162
 CC CONFLICT 1456 1456
 CC CONFLICT 1848 1848
 CC CONFLICT 1892 1892
 CC CONFLICT 1909 1909
 CC CONFLICT 1971 1978
 CC CONFLICT 2050 2050
 CC CONFLICT 2216 2216
 CC CONFLICT 2264 2264
 CC CONFLICT 2289 2289
 CC CONFLICT 2347 2347
 CC CONFLICT 2426 2426
 CC CONFLICT 2609 2610
 CC CONFLICT 2684 2686
 CC CONFLICT 2756 2756
 CC CONFLICT 2897 2898
 CC CONFLICT 3026 3026

FT CONFLICT 3097 3097 N -> F (IN REF. 1).
 FT CONFLICT 3272 3272 S -> A (IN REF. 1).
 FT CONFLICT 3317 3317 S -> R (IN REF. 1).
 FT CONFLICT 3452 3452 S -> Y (IN REF. 1).
 FT CONFLICT 3484 3484 HQMHPPTFTTSH -> DEMSDAGLFTTSE (IN REF. 1).
 FT CONFLICT 3501 3501 PH -> GO (IN REF. 1).
 FT CONFLICT 3508 3508 E -> R (IN REF. 1).
 SQ SEQUENCE 3588 AA; 402072 MW; BC02FBI57D7F1FDB CRC64;
 Query Match 38.4%; Score 48; DB 1; Length 3588;
 Best Local Similarity 45.5%; Pred. No. 92;
 Matches 10; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
 QY 3 GFRRSARRRRPGLDPTPKVLE 24
 DB 1093 GFYSKAGARISGINPTPRVIE 1114
 RESULT 5
 ID TOLQ_PSEAE STANDARD; PRT; 231 AA.
 AC P50598;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE TOLQ protein.
 GN TOLQ OR PA0969.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PAO;
 RX MEDLINE=97113525; PubMed=8955385;
 RA Dennis J.J., Lafontaine E.R., Sokol P.A.;
 RT "Identification and characterization of the tolQ genes of
 Pseudomonas aeruginosa."
 RL J. Bacteriol. 178:7059-7068 (1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PAO1;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
 RA Hickey M.J., Brinkman P.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 opportunistic pathogen."
 RL Nature 406:959-964 (2000).
 CC -!- FUNCTION: Involved in the tonB-independent uptake of proteins (By
 similarity).
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 (Potential).
 CC -!- SIMILARITY: Belongs to the exbB / tolQ family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U39558; AAC44658.1; -.
 CC EMBL; AE004530; BAG04358.1; -.
 CC PIR; C83525; C83525.
 CC InterPro; IPR002898; MotA_ExbB.
 CC Pfam; PF01618; MotA_ExbB; 1.
 CC Transport; Protein transport; Transmembrane; Inner membrane;
 KW

KW Complete proteome. 20 40 POTENTIAL.
 FT TRANSMEM 134 154 POTENTIAL.
 FT TRANSMEM 176 196 POTENTIAL.
 SQ SEQUENCE 231 AA; 25282 MW; 62E22A7E3B83C09D CRC64;
 Query Match 37.6%; Score 47; DB 1; Length 231;
 Best Local Similarity 44.4%; Pred. No. 7.1;
 Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;
 QY 1 KLGFRRSARRRRPGLDPT 18
 DB 89 RAGFKFSRLRQPGVDP 106
 RESULT 6
 ID ORC6_MOUSE STANDARD; PRT; 262 AA.
 AC Q9WUJ8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Origin recognition complex subunit 6.
 GN ORC6L OR ORC6.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Dean F.B., O'Donnell M.;
 RT "cDNA Cloning of a homolog for Saccharomyces cerevisiae ORC6 from Mus
 musculus."
 RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: Component of the origin recognition complex (ORC) that
 binds origins of replication. It has a role in both chromosomal
 replication and mating type transcriptional silencing. Binds to
 the ARS consensus sequence (ACS) of origins of replication in an
 ATP-dependent manner (By similarity).
 CC -!- SUBUNIT: ORC is composed of six subunits (By similarity).
 CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
 CC -!- SIMILARITY: Belongs to the ORC6 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF139659; AAD32667.1; -.
 CC MGD; MGI:1329285; Orc6l.
 CC InterPro; IPR008721; ORC6.
 CC Pfam; PF05460; ORC6; 1.
 CC DNA replication; Nuclear protein; DNA-binding.
 KW SEQUENCE 262 AA; 29188 MW; F8D27BF9C87DE16C CRC64;
 SQ
 Query Match 37.6%; Score 47; DB 1; Length 262;
 Best Local Similarity 40.0%; Pred. No. 8.1;
 Matches 8; Conservative 5; Mismatches 7; Indels 0; Gaps 0;
 QY 5 FRSARRRRPGLDPTPKVLE 24
 DB 205 FSPTLKKKPPGLEPPAKEIE 224
 RESULT 7
 ID RGS3_HUMAN STANDARD; PRT; 519 AA.
 AC P49756; Q8TD59; Q8TD68;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 KW

15-MAR-2004 (Rel. 43, Last annotation update)
Regulator of G-protein signaling 3 (RGS3) (RGP3).
RGS3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX MEDLINE=96178495; PubMed=8602223;
RA Druey K.M., Blumer K.J., Kang V.H., Kehrl J.H.;
RT "Inhibition of G-protein-mediated MAP kinase activation by a new
mammalian gene family.";
RL Nature 379:742-746(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=Brain;
RA Publ H.L. III, Ikeda S.R., Aronstam R.S.;
RT "cDNA clones of human proteins involved in signal transduction
sequenced by the Guthrie cDNA resource center (www.cdna.org).";
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Skuce C.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Inhibits signal transduction by increasing the GTPase
activity of G protein alpha subunits thereby driving them into
their inactive GDP-bound form.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1;
CC IsoId=P49796-1; Sequence=Displayed;
CC Name=2; Synonyms=RGS3f;
CC IsoId=P49796-2; Sequence=VSP_005662;
CC -!- PFM: Phosphorylated by cyclic GMP-dependent protein kinase (By
similarity).
CC -!- SIMILARITY: Contains 1 RGS domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; U27655; AAC50394.1; -;
CC EMBL; AF493927; AAM12641.1; -;
CC EMBL; AF493941; AAM12655.1; -;
CC EMBL; AL162727; CAC78977.1; -;
CC PIR; S78089; S78089.
CC HSP; P49799; IAGR.
CC Gene; HGNC:9999; RGS3.
CC MIM; 602189; -;
CC GO; GO:0005829; Cytosol; TAS.
CC GO; GO:0005096; GTPase activator activity; TAS.
CC GO; GO:000188; P-inactivation of MAPK; TAS.
CC GO; GO:0008277; P-regulation of G-protein coupled receptor pr. .; TAS.
CC InterPro; IPR000342; Regl_Grotein.
CC Pfam; PF00615; RGS; 1.
CC PRINTS; PR01301; RGS-PROTEIN.
CC ProDom; PD001580; Regl_Gprotein; 1.
CC SMART; SM00315; RGS; 1.
CC PROSITE; PS0132; RGS; 1.
CC Signal transduction inhibitor; Alternative splicing; Phosphorylation.
FT DOMAIN 394 510 RGS.
FT VARSLIC 1 313 Missing (in isoform 2).
FT FTIC=VSP_005662.
FT CONFLICT 305 305 K -> R (IN REF. 2).
FT SEQUENCE 519 AA; 56601 MW; F1CFE3F2/D4673A0 CRC64;
Query Match 36.8%; Score 46; DB 1; Length 519;
Best Local Similarity 52.4%; Pred. No. 24;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Matches 11; Conservative 0; Mismatches 8; Indels 2; Gaps 1;
QY 1 KLGFRRSARRRRPGLDTPK 21
DB 355 KLGIFFR--RRNESPGAPGAK 373
RESULT 8
CNAC_RAT
ID CNAC_RAT STANDARD; PRT; 536 AA.
AC P14644;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE CAMP-specific 3',5'-cyclic phosphodiesterase 4C (EC 3.1.4.17)
DE (DPDE1) (Fragment).
GN PDE4C.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95047482; PubMed=7958996;
RA Bolger G.B., Rodgers L., Riggs M.;
RT "Differential CNS expression of alternative mRNA isoforms of the
mammalian genes encoding cAMP-specific phosphodiesterases.";
RL Gene 149:237-244(1994).
RN [2]
RP SEQUENCE OF 153-511 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=89315790; PubMed=2546153;
RA Swinnen J.V., Joseph D.R., Conti M.;
RT "Molecular cloning of rat homologues of the Drosophila melanogaster
dunce cAMP phosphodiesterase: evidence for a family of genes.";
RL Proc Natl Acad Sci U S A. 86:5325-5329(1989).
CC -!- CATALYTIC ACTIVITY: Adenosine 3',5'-cyclic phosphate + H(2)O =
adenosine 5'-phosphate.
CC -!- ENZYME REGULATION: Inhibited by rolipram.
CC -!- PATHWAY: Cyclic nucleotide metabolism.
CC -!- SIMILARITY: Belongs to the cyclic nucleotide phosphodiesterase
family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL; L27061; AAA56858.1; -;
CC EMBL; M25347; AAA41847.1; -;
CC PIR; I67945; I67945.
CC InterPro; IPR003607; Met_phosphohydro.
CC InterPro; IPR002073; PDEase.
CC Pfam; PF00233; PDEase; 1.
CC PRINTS; PR00387; EDIESTERASE1.
CC SMART; SM00471; Hdc; 1.
CC PROSITE; PS00126; PDEASE_I; 1.
CC Hydroxylase; CAMP; Multigene family.
FT NON TER 1 1
FT DOMAIN 524 534 POLY-GLU
FT CONFLICT 218 218 R -> S (IN REF. 2).
FT CONFLICT 507 507 S -> N (IN REF. 2).
FT SEQUENCE 536 AA; 60063 MW; 87D12BE2C4642F3 CRC64;
Query Match 36.8%; Score 46; DB 1; Length 536;
Best Local Similarity 52.9%; Pred. No. 25;
Matches 9; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 4 FRSARRRRPGLDTPK 20
DB 355 KLGIFFR--RRNESPGAPGAK 373

Db 434 FFQGDREERESGLD:SP 450

SYD MYCLE STANDARD; PRT: 589 AA.

RESULT 9

ID SYD MYCLE STANDARD; PRT: 589 AA.

AC P36429; P95671; 36.8%; Score 46; DB 1; Length 589;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Aspartyl-tRNA synthetase (EC 6.1.1.12) (Aspartate--tRNA ligase)

DE (ASPRS) (Antigen T5).

GN ASPS OR MLO501 OR MLCB1259.19.

OS Mycobacterium leprae.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1789;

RN

RP SEQUENCE FROM N.A.

RX MEDLINE=21128732; PubMed=11234002;

RA Cole S.F., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R., Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D., Mungall K., Basham D., Brown D., Chillingworth T., Connor R., Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N., Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S., Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M., Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R., Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R., Barrall B.G.;

RT "Massive gene decay in the leprosy bacillus.";

RL Nature 409:1007-1011(2001).

CC -!- CATALYTIC ACTIVITY: ATP + L-aspartate + tRNA(Asp) = AMP + diphosphate + L-aspartyl-tRNA(Asp).

CC -!- SUBUNIT: Homodimer (By similarity).

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- SIMILARITY: Belongs to class-II aminacyl-tRNA synthetase family.

EMBL: X77655; CAAS4735.1; -.

EMBL: S82268; AAC27132.1; -.

EMBL: AL023591; CAAL9094.1; -.

EMBL: AL583918; CAC30009.1; -.

PIR: S42047; S42047.

HSP: P36419; IEFW.

Leprona; MLO501; -.

HAMAP: MF_00044; -.

InterPro: IPR004524; Asps_bact.

InterPro: IPR004115; GAD_dom.

InterPro: IPR008994; Nucleic_acid_OP.

InterPro: IPR004364; tRNA-synt_2.

InterPro: IPR002312; tRNA-synt_2.

InterPro: IPR004365; tRNA_anti.

InterPro: IPR006195; tRNA_ligase_II.

Pfam: PF02938; GAD; 1.

Pfam: PF00152; tRNA-anti_2; 2.

Pfam: PF01336; tRNA_anti; 1.

PRINTS: PR01042; TRNASYNTHASP.

DR TIGRFAMS; TIGR00459; asps_bact; 1.

DR PROSITE; PS00862; AA_TRNA_LIGASE_II; 1.

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

KW Complete proteome.

SQ SEQUENCE 589 AA; 64630 MW; 6B0EA0675367031C CRC64;

Query Match 36.8%; Score 46; DB 1; Length 589;

Best Local Similarity 44.4%; Pred. No. 27;

Matches 8; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 7 SARRRREPGLDPTPKYLE 24

Db 572 TAOQRKESGIDTPKVE 589

RESULT 10

CN4C_HUMAN STANDARD; PRT: 712 AA.

ID CN4C_HUMAN STANDARD; PRT: 712 AA.

AC Q08493; Q9UN44; Q9UN45; Q9UN46; Q9UPJ6;

DT 01-OCT-1996 (Rel. 34, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE CAMP-specific 3',5'-cyclic phosphodiesterase 4C (EC 3.1.4.17)

DE (DPDE1) (PDE21).

DE PDE4C.

GN Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN

RP SEQUENCE FROM N.A. (ISOFORM PDE4C1).

RX TISSUE=Substantia nigra;

RX MEDLINE=95145731; PubMed=7843419;

RA Engels P., Sullivan M., Mueller T., Luebbert H.;

RT "Molecular cloning and functional expression in yeast of a human CAMP-specific phosphodiesterase subtype (PDE IV-C).";

RL FEBS Lett. 358:305-310(1995).

[2]

RP SEQUENCE FROM N.A. (ISOFORMS PDE4C1; PDE4C2 AND PDE4C3).

RX MEDLINE=20039485; PubMed=10574329;

RA Sullivan M., Olsen A.S., Houslay M.D.;

RT "Genomic organisation of the human cyclic AMP-specific phosphodiesterase PDE4C gene and its chromosomal localisation to 19p13.1, between RAB3A and JUND.";

RL Cell. Signal. 11:735-742(1999).

[3]

RP SEQUENCE FROM N.A.

RA Lamerdin J.E., McCready P.M., Skowronski E., Viswanathan V., Burkhardt-Schultz K., Gordon L., Dias J., Ramirez M., Stilwag S., Phan H., Velasco N., Do L., Regalla W., Terry A., Ganes J., Dangnan L., Poundstone P., Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T., Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A., Saunders C., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S., Carrano A.V.;

RT "Sequence analysis of an ~600 kb region in 19p13.1 between JAK3 and PDE4C.";

RL Submitted (OCT-1998) to the EMBL/GenBank/DBSJ databases.

[4]

RP SEQUENCE OF 452-712 FROM N.A.

RX MEDLINE=94019330; PubMed=8413254;

RA Bolger G., Michaeli T., Martins T., St John T., Steiner B., Rodgers L., Riggs M., Wigler M., Ferguson K.;

RT "A family of human phosphodiesterases homologous to the dunce learning and memory gene product of Drosophila melanogaster are potential targets for antidepressant drugs.";

RL Mol. Cell. Biol. 13:6559-6571(1993).

CC -!- CATALYTIC ACTIVITY: Adenosine 3',5'-cyclic phosphate + H(2)O = adenosine 5'-phosphate.

CC -!- ENZYME REGULATION: Inhibited by rolipram.

CC -!- PATHWAY: Cyclic nucleotide metabolism.

CC -!- ALTERNATIVE PRODUCTS:

CC Event-Alternative splicing, Named isoforms=7;

CC Name=PDE4C1;
 CC IsoId=Q08493-1; Sequence=Displayed;
 CC Name=PDE4C2;
 CC IsoId=Q08493-2; Sequence=VSP_004575;
 CC Name=PDE4C3;
 CC IsoId=Q08493-3; Sequence=VSP_004574;
 CC Name=PDE4C4;
 CC IsoId=Q08493-4; Sequence=Not described;
 CC Name=PDE4C5;
 CC IsoId=Q08493-5; Sequence=Not described;
 CC Name=PDE4C6;
 CC IsoId=Q08493-6; Sequence=Not described;
 CC Name=PDE4C7;
 CC IsoId=Q08493-7; Sequence=Not described;
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN VARIOUS TISSUES BUT NOT IN CELLS
 CC OF THE IMMUNE SYSTEM.
 CC -1- SIMILARITY: Belongs to the cyclic nucleotide phosphodiesterase
 CC family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Z46632; CAA86601.1; -
 CC EMBL; AF157816; AAD47053.1; -
 CC EMBL; AF157811; AAD47053.1; JOINED.
 CC EMBL; AF157814; AAD47053.1; JOINED.
 CC EMBL; AF157815; AAD47053.1; JOINED.
 CC EMBL; AF157816; AAD47054.1; -
 CC EMBL; AF157812; AAD47054.1; JOINED.
 CC EMBL; AF157814; AAD47054.1; JOINED.
 CC EMBL; AF157815; AAD47054.1; JOINED.
 CC EMBL; AF157816; AAD47055.1; -
 CC EMBL; AF157814; AAD47055.1; JOINED.
 CC EMBL; AF157815; AAD47055.1; JOINED.
 CC EMBL; AC005759; AAC83047.1; -
 CC EMBL; L20968; AAA03591.1; -
 CC PIR; S71626; S71626
 CC PDB; 1LXU; 26-JUN-02
 CC Genew; HGNC:18782; PDE4C.
 CC MIM; 600128; -
 CC GO; GO:0004115; F:CAMP-specific phosphodiesterase activity; TAS.
 CC InterPro; IPR003607; Met phosphohydro.
 CC InterPro; IPR002073; PDEase.
 CC Pfam; PF00233; PDEase; 1
 CC PRINTS; PR00387; PDIESTERASE1.
 CC SMART; SM00471; Hdc; 1.
 CC PROSITE; PS00126; PDEASE_I; 1.
 CC Hydrolase; CAMP; Multigene family.
 CC VARSPLIC 1 81
 CC MENLGVGDGAEACSLRSRGHSMTRAPKHLWRQPREPR
 CC IQQFVSDPKSAGCERDLSPRELKSLSPVSSCR
 CC -3- MCGPPAPVPFGSPRSGSPGLFRKLNVQSIRL
 CC QRFVVARPLC (in isoform PDE4C3).
 CC -----
 CC VARSPLIC 1 106
 CC Missing (in isoform PDE4C2).
 CC /FTIG-VSP 004574.
 CC K -> N (IN REF. 2).
 CC D -> Y (IN REF. 2).
 CC EL -> DV (IN REF. 1).
 CC NSE -> K (IN REF. 3).
 CC EL -> DV (IN REF. 1).
 CC SEQUENCE 712 AA; 79901 MW; 1932116C9CE0322C CRC64;
 CC -----
 CC Query Match 36.8%; Score 46; DB 1; Length 712;
 CC Best Local Similarity 52.9%; Pred. No. 33;
 CC Matches 9; Conservative 5; Indels 0; Gaps 0;
 CC 4 PFRSRRRREPGLDTP 20
 CC ||: ||| ||| :|

DB 568 FFOQDRERESGLDISP 584
 RESULT 11
 ID DP2L_HALN1 STANDARD; PRT; 1370 AA.
 AC Q9HMX8;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA polymerase II large subunit (EC 2.7.7.7) (Pol II) [Contains: Hsp-
 DE NRC1 polC intein (Hsp-NRC1 pol2 intein)].
 GN POLC OR POLA2 OR VNG2338G.
 OS Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).
 OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950;
 RA NG W.V., Kennedy S.P., Mahairas G.G., Berquist B., Pan M.,
 RA Shukla H.D., Laskey S.R., Baliga N.S., Thorsson V., Sbrogna J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,
 RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
 RA Isenbarger T.A., Beck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dasarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 CC -1- FUNCTION: Possesses two activities: a DNA synthesis (polymerase)
 CC and an exonucleolytic activity that degrades single stranded DNA
 CC in the 3' to 5' direction. Has a template-primer preference which
 CC is characteristic of a replicative DNA polymerase (by similarity).
 CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate = N diphosphate
 CC + {DNA}(N).
 CC -1- CATALYTIC ACTIVITY: Degradation of single-stranded DNA. It acts
 CC progressively in a 3'- to 5'-direction, releasing nucleoside 5'-
 CC phosphates.
 CC -1- SUBUNIT: Heterodimer of a large subunit and a small subunit (By
 CC similarity).
 CC -1- PTM: This protein undergoes a protein self splicing that involves
 CC a post-translational excision of the intervening region (intein)
 CC followed by peptide ligation (potential).
 CC -1- SIMILARITY: Belongs to the archaeal DNA polymerase II family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AG005116; AAC20443.1; -
 CC PIR; G84384; G84384.
 CC HAMAP; MF_00324; -; 1.
 CC InterPro; IPR003587; Hedgehog hint N.
 CC InterPro; IPR003586; Hedgehog hint C.
 CC InterPro; IPR006141; Intein S.
 CC InterPro; IPR004475; PolC DP2.
 CC Pfam; PF03833; PolC_DP2; 1.
 CC SMART; SM00305; HintC; 1.
 CC SMART; SM00306; HintN; 1.
 CC TIGRFAMs; TIGR01443; Intein Cterm; 1.
 CC TIGRFAMs; TIGR01445; Intein Nterm; 1.
 CC TIGRFAMs; TIGR00354; PolC; 1.
 CC PROSITE; PS50818; INTEIN_C_TER; FALSE_NEG.
 CC PROSITE; PS50817; INTEIN_N_TER; 1.
 CC Transfaser; DNA-directed DNA polymerase; DNA replication; Hydrolase;
 CC Nuclease; Exonuclease; DNA-binding; Multifunctional enzyme;
 CC Autocatalytic cleavage; Protein binding; Complete proteome.
 CC CHAIN 1 925 DNA POLYMERASE II LARGE SUBUNIT, 1ST PART
 FT

```
FT CHAIN 926 1120 (POTENTIAL).
FT HSP-NRC1 POLC INTIN (POTENTIAL).
FT CHAIN 1121 1370 DNA POLYMERASE II LARGE SUBUNIT, 2ND PART
FT CHAIN (POTENTIAL).
SQ SEQUENCE 1370 AA; 150295 MW; 07878AA9976790C9 CRC64;

Query Match 36.8%; Score 46; DB 1; Length 1370;
Best Local Similarity 55.6%; Pred. No. 66;
Matches 10; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 5 FRSARRRRPGLDPTPKV 22
DB 20 FEVAARARRRRGDDPTFNV 37

RESULT 12
SOC6_HUMAN
ID SOC6_HUMAN STANDARD; PRT; 485 AA.
AC O1452;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Suppressor of cytokine signaling 6 (SOCS-6) (Nck, Ash and
DE phospholipase C gamma-binding protein) (Nck-associated protein 4)
DE (NAP-4) (Fragment).
DE SOCS6 OR NAP4.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98008866; PubMed=9344857;
RA Matuoka K., Miki H., Takahashi K., Takenawa T.;
RT "A novel ligand for an SH3 domain of the adaptor protein Nck bears an
RT SH2 domain and nuclear signaling motifs."
RL Biochem. Biophys. Res. Commun. 239:488-492(1997).
CC -!- FUNCTION: SOCS family proteins form part of a classical negative
CC feedback system that regulates cytokine signal transduction.
CC -!- SUBUNIT: Interacts, via the third proline-rich region, with the
CC second SH3 domain of the adaptor protein Nck. Also interacts with
CC Grb2 and phospholipase C-gamma.
CC -!- TISSUE SPECIFICITY: Expressed in brain and leukocytes. Also in
CC fetal lung fibroblasts and fetal brain.
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -!- SIMILARITY: Contains 1 SOCS box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
EMBL; AB005216; BAA22432.1; .
PIR; PC4427; PC4427.
HSSP; P23727; 2PNB.
CO; GO:0017124; F:SH3-domain binding; NAS.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001496; SOCS_C.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR SMART; SM00253; SOCS; 1.
DR PROSITE; PS50001; SH2; 1.
DR PROSITE; PS50225; SOCS; 1.
KW SH2 domain; Growth regulation; Signal transduction inhibitor.
FT DOMAIN 82 202 PRO-RICH.
FT DOMAIN 301 381 PRO-RICH.
FT DOMAIN 398 507 SH2.
FT DOMAIN 502 552 SOCS BOX.
FT DOMAIN 84 97 POLY-PRO.
FT DOMAIN 141 149 POLY-GLY.
FT DOMAIN 181 185 POLY-GLN.
FT DOMAIN 186 195 POLY-PRO.
FT DOMAIN 301 310 POLY-PRO.
FT DOMAIN 341 348 POLY-PRO.
SQ SEQUENCE 579 AA; 62783 MW; FAB66BF2A0BE685A CRC64;

Query Match 36.0%; Score 45; DB 1; Length 579;
Best Local Similarity 55.0%; Pred. No. 38;
Matches 11; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRRPGLDPTP 20
DB 360 KNGKFLYFLRSRVFGLPPTP 379

RESULT 13
SOC6_MOUSE
ID SOC6_MOUSE STANDARD; PRT; 579 AA.
AC Q8VHQ2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Suppressor of cytokine signaling 6 (SOCS-6).
DE SOCS6 OR CISH7.
GN Mus musculus (Mouse).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RA Hilton D.J., Viney E.M., Alexander W.S., Willson T.A., Nicola N.A.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: SOCS family proteins form part of a classical negative
CC feedback system that regulates cytokine signal transduction.
CC -!- SIMILARITY: Contains 1 SH2 domain.
CC -!- SIMILARITY: Contains 1 SOCS box domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 12:51:12 ; Search time 3.62011 Seconds
(without alignments)
2091.769 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1188
Perfect score: 125
Sequence: 1 KLGFRRSARRRRPGLDPTPKVLE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: sp_archaea.*
 - 2: sp_bacteria.*
 - 3: sp_fungi.*
 - 4: sp_human.*
 - 5: sp_invertebrate.*
 - 6: sp_mammal.*
 - 7: sp_mhc.*
 - 8: sp_organelle.*
 - 9: sp_phase.*
 - 10: sp_plant.*
 - 11: sp_rodent.*
 - 12: sp_virus.*
 - 13: sp_vertebrate.*
 - 14: sp_unclassified.*
 - 15: sp_rvirus.*
 - 16: sp_bacteriapi.*
 - 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	823	4 Q8WY18	Q8WY18 homo sapien
2	93	74.4	823	11 Q8CEB4	Q8CEB4 mus musculus
3	93	74.4	1188	11 Q7TQC3	Q7TQC3 mus musculus
4	54	43.2	958	11 Q8BWF5	Q8BWF5 mus musculus
5	50.5	40.4	59	13 Q804T8	Q804T8 fugu rubrip
6	50	40.0	179	16 Q87ZC4	Q87ZC4 pseudomonas
7	49	39.2	189	4 Q96X04	Q96X04 homo sapien
8	49	39.2	216	2 Q8VPO5	Q8VPO5 micrococccus
9	49	39.2	241	2 Q7WTF7	Q7WTF7 streptomyce
10	49	39.2	309	5 Q9GRD8	Q9GRD8 drosophila
11	49	39.2	472	16 Q9RX81	Q9RX81 streptomyce
12	49	39.2	740	4 Q8IZ41	Q8IZ41 homo sapien
13	49	39.2	947	5 Q8SYA1	Q8SYA1 drosophila
14	49	39.2	947	5 Q8MLT8	Q8MLT8 drosophila
15	49	39.2	1218	5 Q9V8R6	Q9V8R6 drosophila
16	48	38.4	168	13 Q7SVI3	Q7SVI3 gallus gall

17	48	38.4	283	13 Q7SVI2	Q7SVI2 gallus gall
18	48	38.4	373	16 Q66737	Q66737 aquifex aeo
19	48	38.4	408	13 Q8YI11	Q8YI11 gallus gall
20	48	38.4	441	13 Q7SVI0	Q7SVI0 gallus gall
21	48	38.4	528	3 P87239	P87239 schizosacch
22	48	38.4	798	3 Q14026	Q14026 schizosacch
23	48	38.4	799	13 Q7SYH9	Q7SYH9 gallus gall
24	47	37.6	66	16 Q7UPB4	Q7UPB4 rhodopirell
25	47	37.6	204	10 Q94CP9	Q94CP9 oryza sativ
26	47	37.6	231	2 Q9WXX3	Q9WXX3 pseudomonas
27	47	37.6	231	16 Q88N18	Q88N18 pseudomonas
28	47	37.6	237	4 Q8N0T0	Q8N0T0 homo sapien
29	47	37.6	248	10 Q8LN93	Q8LN93 oryza sativ
30	47	37.6	300	16 Q8XUM2	Q8XUM2 raistonia s
31	47	37.6	2144	2 Q8G982	Q8G982 planktothri
32	46.5	37.2	403	16 Q82AM6	Q82AM6 streptomyce
33	46	36.8	120	16 Q8ZL06	Q8ZL06 salmonella
34	46	36.8	120	16 Q8ZL06	Q8ZL06 salmonella
35	46	36.8	132	4 Q8WV02	Q8WV02 homo sapien
36	46	36.8	284	4 Q96NV5	Q96NV5 homo sapien
37	46	36.8	319	4 Q8NFN6	Q8NFN6 homo sapien
38	46	36.8	426	4 Q9UPJ5	Q9UPJ5 homo sapien
39	46	36.8	427	4 Q76104	Q76104 homo sapien
40	46	36.8	427	4 Q43851	Q43851 homo sapien
41	46	36.8	518	4 Q43850	Q43850 homo sapien
42	46	36.8	700	4 P78505	P78505 homo sapien
43	46	36.8	708	16 Q92YMS	Q92YMS rhizobium m
44	46	36.8	782	4 Q76105	Q76105 homo sapien
45	46	36.8	791	4 Q43849	Q43849 homo sapien

ALIGNMENTS

RESULT 1

Q8WY18	PRELIMINARY;	PRT;	823 AA.
ID	Q8WY18		
AC	Q8WY18;		
DT	01-MAR-2002 (TREMBLrel. 20, Created)		
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)		
DE	01-JUN-2003 (TREMBLrel. 24, Last annotation update)		
DE	MSTP018.		
GN	MSTC18.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
OX	NCBI TaxID=9606;		
[1]			
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Adcta;		
RA	Xu Y.Y., Sun L.Z., Wu Q.Y., Liu Y.Q., Liu B., Zhao B., Wang X.Y.,		
RA	Song L., Ye J., Sheng H., Gao Y., Zhang C.L., Zhang J., Wei Y.J.,		
RA	Sun Y.H., Jiang Y.X., Zhao X.W., Liu S., Liu L.S., Ding J.F.,		
RA	Gao R.L., Qiang B.Q., Yuan J.G., Liew C.C., Zhao M.S., Hui R.T.;		
RL	Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF111799; AAL39001.1;		
DR	GO; GO:0008305; C:integrin complex; IEA.		
DR	GO; GO:0004895; F:cell adhesion receptor activity; IEA.		
DR	GO; GO:0007160; P:cell-matrix adhesion; IEA.		
DR	InterPro; IPR000413; Integrin_alpha.		
DR	Pfam; PF01839; FG-GAP; 3.		
DR	PRINTS; PR01185; INTEGRINA.		
DR	SMART; SM00191; Int. alpha; 4.		
SQ	SEQUENCE 823 AA; 92672 NW; DE4E78079DCD4925 CRC64;		

Query Match 100.0%; Score 125; DB 4; Length 823;
Best Local Similarity 100.0%; Pred. No. 1.le-10;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRRSARRRRPGLDPTPKVLE 24

DB 800 KLGFRRSARRRRPGLDPTPKVLE 823

```

RESULT 2
Q8CE84 PRELIMINARY; PRT; 823 AA.
ID AC Q8CE84;
AC Q8CE84;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE MSTP018 homolog.
GN 4732459H24RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK028821; BAC26137.1; -.
DR MGD; MGI:2442114; 4732459H24RIK.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR PRINTS; PR01185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 4.
SQ SEQUENCE 823 AA; 92264 MW; A330236324A0E089 CRC64;

Query Match 74.4%; Score 93; DB 11; Length 823;
Best Local Similarity 70.8%; Pred. No. 1e-05;
Matches 17; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREPGLDPTPKVLE 24
DB 800 KLGFFRSARRRREPGLDPTPKVLE 823

RESULT 3
Q7TQC3 PRELIMINARY; PRT; 1188 AA.
ID AC Q7TQC3;
AC Q7TQC3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE All integrin.
GN ITGall.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg E.,
RC Guilberg D.;
RT "allb1 integrin is important for mesenchymal cell function:
RT elimination of allb1 leads to dwarfism."
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC Johansson M., Popova S.N.;
RC Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV124460; ANM62130.1; -.
KW Integrin.
SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75E4271 CRC64;

Query Match 74.4%; Score 93; DB 11; Length 1188;
Best Local Similarity 70.8%; Pred. No. 1.5e-05;

```

```

Matches 17; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREPGLDPTPKVLE 24
DB 1165 KLGFFRSARRRREPGLDPTPKVLE 1188

RESULT 4
Q8BMF5 PRELIMINARY; PRT; 956 AA.
ID AC Q8BMF5;
AC Q8BMF5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Glutamate receptor.
GN GRIK4 OR 6330551X01RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Medulla oblongata;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK032029; BAC27860.1; -.
DR MGD; MGI:95817; Grik4.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005234; F:glutamate-gated ion channel activity; IEA.
DR GO; GO:0004970; F:ionotropic glutamate receptor activity; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0005267; F:potassium channel activity; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0006813; P:potassium ion transport; IEA.
DR InterPro; IPR01828; ANF_receptor.
DR InterPro; IPR001320; Ion_glu_receptor.
DR InterPro; IPR001622; K_channel_pore.
DR InterPro; IPR001508; NMDA_receptor.
DR InterPro; IPR001311; SBP/Glu_receptor.
DR Pfam; PF01094; ANF_receptor; 1.
DR Pfam; PF00060; Lig_chan; 1.
DR PRINTS; PR00177; NMDARECEPTOR.
DR SMART; SM00079; PBpe; 1.
SQ SEQUENCE 956 AA; 107299 MW; 4A1E39518FBA7A38 CRC64;

Query Match 43.2%; Score 54; DB 11; Length 956;
Best Local Similarity 68.8%; Pred. No. 13;
Matches 11; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 9 RRRRREPGLDPTPKVLE 24
DB 862 RRRRREPGLDPTPKVLE 877

RESULT 5
Q804T8 PRELIMINARY; PRT; 59 AA.
ID AC Q804T8;
AC Q804T8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bone Gla protein.
GN BGP.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Perciformes; Tetraodontiformes;

```



```

OC Tetradontidae; Tetradontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RA Laize V., Canela M.L.;
RT "Identification of Takifugu rubripes osteocalcin (BGP) by comparative
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF478914; AAC24898.1; -.
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR InterPro; IPR002384; GLA_bone.
DR InterPro; IPR000294; VitK_ggp_GLA.
DR Pfam; PF00594; Gla; 1.
DR PRINTS; PRO0002; GLABONE.
DR PROSITE; PS00011; GLU CARBOXYLATION; 1.
SQ SEQUENCE 59 AA; 6506 MW; 97112FFDA6A6AE1DB CRC64;

Query Match 40.4%; Score 50.5; DB 13; Length 59;
Best Local Similarity 68.8%; Pred. No. 2.6;
Matches 11; Conservative 2; Mismatches 2; Indels 1; Gaps 1;

QY 9 RRRRREPGLDTPKYLE 24
    ||||| ||| :|||
Db 11 RRRRAPG-EPTPQLE 25

RESULT 6
Q87ZC4 PRELIMINARY; PRT; 179 AA.
AC Q87ZC4;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Cytochrome b561, putative.
GN "P8PTQ3505.
OS Pseudomonas syringae (pv. tomato).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=323;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DC3000;
RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
RA Barry K., Uterback T., Van Aken S., Feldblyum T., Gwinn M.,
RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidsen T.,
RA White O., Fraser C., Collmer A.;
RL "Complete sequence of Pseudomonas syringae.";
RT Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016868; AAC056980.1; -.
DR TIGR; PSPT03505; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0016491; F:oxidoreductase activity; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000516; Ni_hydr_CyTB.
DR Pfam; PF01292; Ni_hydr_CyTB; 1.
KW Complete proteome.
SQ SEQUENCE 179 AA; 20500 MW; 0F915D51AF24F92F CRC64;

Query Match 40.0%; Score 50; DB 16; Length 179;
Best Local Similarity 57.9%; Pred. No. 9.8;
Matches 11; Conservative 2; Mismatches 4; Indels 2; Gaps 1;

QY 2 LGFRSARRRRR--PGLDP 18
    ||||| ||| :|||
Db 61 LAFFVARRRRRPPGITP 79

RESULT 7
Q96N04 PRELIMINARY; PRT; 189 AA.
AC Q96N04;

```

```

DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein FLJ31614.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T., Kodaira H.,
RA Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kanihara K.,
RA Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,
RA Osuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J., Isono Y.,
RA Kawai-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Murakawa K.,
RA Sukehori K., Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B.,
RA Suzuki Y., Sugano S., Nagahari K., Masuho Y., Nagai K., Isogai T.;
RL "NEDO human cDNA sequencing project.";
RT Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK056176; BAB71112.1; -.
DR GO; GO:0005509; P:calcium ion binding; IEA.
DR InterPro; IPR002048; EF-hand.
DR Pfam; PF00036; efhand; 2.
DR SMART; SM00054; EFH; 2.
DR PROSITE; PS00018; EF_HAND; 2.
KW Hypothetical protein.
SQ SEQUENCE 189 AA; 20736 MW; 472825570664F33D CRC64;

Query Match 39.2%; Score 49; DB 4; Length 189;
Best Local Similarity 52.2%; Pred. No. 15;
Matches 12; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 LGFRSARRRRRPGLDTPKYLE 24
    ||||| ||| :|||
Db 72 LGSIRGGRRDGMGLDPAPVSE 94

RESULT 8
Q8VPQ5 PRELIMINARY; PRT; 216 AA.
AC Q8VPQ5;
DT 01-MAR-2002 (T-EMBLrel. 20, Created)
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE MC16.
OS Micrococcus sp. 28.
OG Plasmid pSD10.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococcaceae; Micrococcaceae; Micrococcus.
OX NCBI_TaxID=161213;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=28;
RA Zhong Z., Caspi R., Mincer T., Helinski D., Knauf V., Boardman K.,
RA Wilkinson J.E., Shea T., DeLoughery C., Toukdarian A.;
RT "A 50 kb plasmid rich in mobile gene sequences isolated from a marine
RT Micrococcus.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY034092; AAK62490.1; -.
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
KW Plasmid.
SQ SEQUENCE 216 AA; 22533 MW; 7A082298C81297AF CRC64;

Query Match 39.2%; Score 49; DB 2; Length 216;
Best Local Similarity 47.4%; Pred. No. 17;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 6 RSARRRRRPGLDTPKYLE 24
    ||||| ||| :|||
Db 165 RFSRRRRRPGVRRPRDL 183

```


DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Similar to hypothetical protein FLJ31614.
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Muscle;
 RA Strausberg R.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RM EMBL: BC023566; AAH23566.1; -
 DR F1R; PT0272; PT0271.
 DR GO: GO:000509; F:Calcium ion binding; IEA.
 DR GO: GO:000525; F:GTP binding; IEA.
 DR GO: GO:0003928; F:RAB small monomeric GTPase activity; IEA.
 DR GO: GO:0003930; F:RAS small monomeric GTPase activity; IEA.
 DR GO: GO:0003931; F:Rho small monomeric GTPase activity; IEA.
 DR GO: GO:0006886; P:intracellular protein transport; IEA.
 DR GO: GO:0007264; P:small GTPase mediated signal transduction; IEA.
 DR InterPro: IPR002048; EF-hand.
 DR InterPro: IPR003579; GTPase_Rab.
 DR InterPro: IPR003577; GTPase_Ras.
 DR InterPro: IPR003578; GTPase_Rho.
 DR InterPro: IPR002041; RAN.
 DR InterPro: IPR001806; Ras trnsfrmg.
 DR InterPro: IPR005225; Small_GTP.
 DR Pfam: PF00071; ras; 1.
 DR PRINTS: PR00449; RASTRNSFRMG.
 DR SMART: SM00054; EPH; 2.
 DR SMART: SM00175; RAB; 1.
 DR SMART: SM00176; RAN; 1.
 DR SMART: SM00173; RAS; 1.
 DR SMART: SM00174; RHO; 1.
 DR TIGRFAMs: TIGR00231; small_GTP; 1.
 DR PROSITE: PS00018; EF_HAND; 2.
 KW Hypothetical protein.
 SQ SEQUENCE 740 AA; 82879 MW; 89AFCF4C159760F0 CRC64;

Query Match 39.2%; Score 49; DB 4; Length 740;
 Best Local Similarity 52.2%; Pred. No. 59;
 Matches 12; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 2 LGFFRSARRRRREFGLDPTPKVLE 24
 DB 72 LGSILGRRRDWGLDPAPAVSE 94

RESULT 13

ID Q8SYA1 PRELIMINARY; PRT; 947 AA.
 AC Q8SYA1
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE RH10407p.
 GN CG7097.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Faise B.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Celniker S.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY071688; AAL49310.1; -
 DR FlyBase; FBgn0034421; CG7097.
 DR GO: GO:0005524; F:ATP binding; IEA.
 DR GO: GO:0004574; F:protein serine/threonine kinase activity; IEA.
 DR GO: GO:0005083; F:small GTPase regulatory/interacting protein.; IEA.
 DR GO: GO:0016740; F:transferase activity; IEA.
 DR GO: GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro: IPR001180; Citron.
 DR InterPro: IPR000719; Prot_kinase.
 DR InterPro: IPR002290; Ser_Chtr_kinase.
 DR Pfam: PF00780; CNH; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR ProDom: PD000001; Prot_kinase; 1.
 DR SMART: SM00036; CNH; 1.
 DR SMART: SM00220; S_TKG; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
 DR ATP-binding; Transferase.
 KW ATP-binding; Transferase.
 SQ SEQUENCE 947 AA; 105224 MW; 814262CDB95FF56D CRC64;
 Query Match 39.2%; Score 49; DB 5; Length 947;
 Best Local Similarity 52.2%; Pred. No. 76;
 Matches 12; Conservative 1; Mismatches 4; Indels 6; Gaps 1;
 QY 6 RSARRRREP-----GLDPTPKV 22
 DB 552 RSHKRRTTPRPISNGLPPTPKV 574
 RESULT 14
 ID Q8MLI8
 AC Q8MLI8; PRELIMINARY; PRT; 947 AA.
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CG7097-PB.
 GN CG7097.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RM MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier A.S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iqbalwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M., Wang X.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banzon J., An H., Baldwin D., Banton J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwan C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RA "Sequencing of *Drosophila melanogaster* genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Masra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RA "Annotation of *Drosophila melanogaster* genome."
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RA Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RL EMBL; AE003796; AM70845.1; ..
DR FlyBase; FBgn0034421; CG7097.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . ; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002230; Ser_Thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00780; CNH; 1.
DR ProDom; PD00069; pkinase; 1.
DR SMART; SM00036; CNH; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TyKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS0011; PROTEIN_KINASE_DOM; 1.
DR ATP-binding; Transferrase.
SQ SEQUENCE 947 AA; 105195 MW; 480EC84EABE0D0E0 CRC64;

Query Match

39.2%; Score 49; DB 5; Length 947;

Best Local Similarity 52.2%; Pred. No. 76;
Matches 12; Conservative 1; Mismatches 4; Indels 6; Gaps 1;
QY 6 RSARRRRREP-----GLDPTPKV 22
DB 552 RSHKRHTPPRPISGNGLPPTPKV 574
RESULT 15
QV9VR6
ID QV9VR6 PRELIMINARY; PRT; 1218 AA.
AC QV9VR6;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE CG7097 protein.
GN CG7097.

OS *Drosophila melanogaster* (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]_taxID=7227;
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Fertaz C., Ferreira S., Fleischmann W.,
RA Foele C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
RL EMBL; AE003796; AM75595.1; ..
DR HSP; Q63450; IAO6.
DR FlyBase; FBgn0034421; CG7097.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004574; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0005083; F:small GTPase regulatory/interacting protein. . ; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR001180; Citron.

Tue Sep 21 14:48:45 2004

DR InterPro: IPR000719; Prot_kinase.

DR InterPro: IPR002290; Ser_Thr_kinase.

DR Pfam: PF00780; CNH; 1.

DR Pfam: PF00069; Pkinase; 1.

DR ProDom: PD000001; Prot_kinase; 1.

DR SMART: SM00036; CNH; 1.

DR SMART: SM00220; S_TKC; 1.

DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.

DR PROSITE: PS00111; PROTEIN_KINASE_DOM; 1.

KW ATP-binding; Transferase.

SQ SEQUENCE 1218 AA; 132395 MW; 78A7FAFF880CEDC8 CRC64;

Query Match 39.2%; Score 49; DB 5; Length 1218;

Best Local Similarity 52.2%; Pred. No. 95;

Matches 12; Conservative 1; Mismatches 4; Indels 6; Gaps 1;

QY 6 RSARRREP-----GLDPTPKV 22

Db 823 RSHKRRHPPRFISNGLPPTPKV 845

Search completed: September 21, 2004, 13:03:45

Job time : 6.62011 secs

This page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 12:43:42 ; Search time 4.57542 Seconds
(without alignments)
1482.081 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1188

Perfect score: 125
Sequence: 1 KLGFFRSARRRRFGLDTPKYLE 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	125	100.0	24	AAB30927	Aab30927 Peptide d
2	125	100.0	1120	ABR58365	AbR58365 Human NOV
3	125	100.0	1188	ABR30929	AbR30929 Amino aci
4	125	100.0	1188	AAU14467	Aau14467 Human nov
5	125	100.0	1188	AAU14231	Aau14231 Human nov
6	125	100.0	1188	AAU14231	AAU14231 Human nov
7	125	100.0	1188	AAU10551	Aau10551 Human A25
8	125	100.0	1188	7 ADE09956	AdE09956 Novel pro
9	125	100.0	1189	3 AAB25582	AbA25582 ITGAl1 pr
10	125	100.0	1189	4 ABG12949	AbG12949 Novel hum
11	125	100.0	1189	6 ABR58364	AbR58364 Human NOV
12	125	100.0	1189	6 ADA27054	Ada27054 Human nov
13	125	100.0	1189	7 ADE63570	AdE63570 Human pro
14	125	100.0	1189	8 ADE86584	AdE86584 Novel hum
15	93	74.4	1188	4 AAB50087	AbA50087 Murine A2
16	93	74.4	1188	5 AAU10552	Aau10552 Murine pr
17	92	73.6	545	5 ABB72300	AbB72300 Rat prote
18	92	73.6	688	5 ABB72300	AbB72300 Rat prote
19	75	60.0	15	4 AAB30928	AbA30928 Rat Prote
20	52	41.6	956	7 ADE62089	AdE62089 Rat Prote
21	51	40.8	161	4 AAM96433	Aam96433 Human rep
22	51	40.8	161	4 ABB10951	AbB10951 Human ova
23	50	40.0	1024	6 ABB14615	AbB14615 Novel hum
24	49	39.2	189	6 ADA54842	Ada54842 Human pro
25	49	39.2	466	4 AAU58625	Aau58625 Propionib

26	49	39.2	466	6	ABM55144	Abm55144 Propionib
27	49	39.2	616	6	ABM65406	Abm65406 Propionib
28	49	39.2	832	3	AAU58663	Aau58663 Human tyr
29	49	39.2	832	5	AAU71093	Aau71093 Human can
30	49	39.2	1218	4	ABR62736	AbR62736 Drosophil
31	48	38.4	100	4	AAO03386	Aao03386 Human pol
32	48	38.4	373	2	AAW24249	Aaw24249 Aquifex a
33	48	38.4	373	2	ABU57352	Abu57352 Aquifex a
34	48	38.4	3588	2	AAU34712	Aau34712 Bacillus
35	47	37.6	69	4	AAU58901	Aau58901 Propionib
36	47	37.6	69	4	AAU58844	Aau58844 Propionib
37	47	37.6	69	6	ABM62363	Abm62363 Propionib
38	47	37.6	69	6	ABM55420	Abm55420 Propionib
39	47	37.6	72	2	AAW27427	Aaw27427 Human CRA
40	47	37.6	614	4	ABG24990	AbG24990 Novel hum
41	47	37.6	964	6	ABU52612	Abu52612 Human NOV
42	47	37.6	1245	4	ABG25469	AbG25469 Novel hum
43	46	36.8	99	4	AAU40886	Aau40886 Propionib
44	46	36.8	99	6	ABM37405	Abm37405 Propionib
45	46	36.8	114	4	AAU57162	Aau57162 Propionib

ALIGNMENTS

RESULT 1
AAB30927
ID AAB30927 standard; peptide; 24 AA.
XX
AC AAB30927;
XX
DT 02-APR-2001 (first entry)
DE Peptide derived from a human alpha11 integrin chain.
XX
KW Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;
KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;
KW osteoporosis; cartilage damage; bone damage; cartilage.
XX
OS Homo sapiens.
XX
PN WO200075187-A1.
XX
PD 14-DEC-2000.
XX
PF 31-MAY-2000; 2000WO-SE001135.
XX
PR 03-JUN-1999; 99SE-00002056.
XX
XX (ACTI-) ACTIVE BIOTECH AB.
PI Gullberg D;
XX
DR WPI; 2001-071061/08.
XX
PT Integrin subunit alpha 11 or integrin heterodimer comprising subunit
PT alpha 11 in association with subunit beta, useful for treating muscle
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.
XX
PS Claim 23; Page 40; 79pp; English.

CC The present sequence is derived from the cytoplasmic domain of the human
CC integrin subunit, designated alpha11. The alpha11 polynucleotide and
CC polypeptide are useful as markers of cell target molecules, such as
CC fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally
CC derived cells or stem cells. They are also used for determining the
CC differential stage of cells during differentiation, development in
CC pathological conditions in tissue regeneration, in transplantation or in
CC therapeutic and physiological repair of tissues. The pathological
CC conditions involving subunit alpha11 are selected from damage of cells,
CC muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,
CC osteoarthritis and osteoporosis, damage of cartilage and bone, and

CC cartilage and bone diseases. The polypeptide is useful for detecting the
 CC formation of cartilage during embryonic development, for detecting
 CC physiological therapeutic repair of cartilage and muscle, for selection
 CC and analysis, or for sorting, isolating or purification of chondrocytes
 CC and muscle cells, for detecting regeneration of cartilage or chondrocytes
 CC during transplantation of cartilage or chondrocytes during
 CC transplantation of cartilage or chondrocytes, respectively, or of muscle
 CC or muscle cells during transplantation of muscle or muscle cells,
 CC respectively, and for studies of differentiation or chondrocytes or
 CC muscle cells

XX Sequence 24 AA;

Query Match 100.0%; Score 125; DB 4; Length 24;
 Best Local Similarity 100.0%; Pred. No. 1.3e-10;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRRSARRRRRPGDPTPKVLE 24
 |||||
 Db 1 KLGFRRSARRRRRPGDPTPKVLE 24
 |||||

RESULT 2

ID ABR58365
 AC ABR58365 standard; protein; 1120 AA.

XX ABR58365;

DT 07-JUN-2003 (first entry)

XX Human NOV2b.

XX Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;
 KW antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;
 KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;
 KW neurodegenerative disorder; Alzheimer's disease; immune disorder;
 KW haematopoietic disorder.

XX Homo sapiens.

XX WO2003029423-A2.

XX 10-APR-2003.

XX 02-OCT-2002; 2002WO-US031358.

XX 02-OCT-2001; 2001US-0326483P.
 PR 05-OCT-2001; 2001US-0327342P.
 PR 09-OCT-2001; 2001US-0327917P.
 PR 09-OCT-2001; 2001US-0328029P.
 PR 09-OCT-2001; 2001US-0328044P.
 PR 09-OCT-2001; 2001US-0328056P.
 PR 12-OCT-2001; 2001US-0328849P.
 PR 15-OCT-2001; 2001US-0329414P.
 PR 17-OCT-2001; 2001US-0330142P.
 PR 22-OCT-2001; 2001US-0341058P.
 PR 24-OCT-2001; 2001US-0339266P.
 PR 29-OCT-2001; 2001US-0343629P.
 PR 01-NOV-2001; 2001US-0349575P.
 PR 12-APR-2002; 2002US-0346357P.
 PR 12-APR-2002; 2002US-0371972P.
 PR 17-APR-2002; 2002US-0371980P.
 PR 19-APR-2002; 2002US-0373261P.
 PR 23-APR-2002; 2002US-0373805P.
 PR 16-MAY-2002; 2002US-0381101P.
 PR 17-MAY-2002; 2002US-0381635P.
 PR 29-MAY-2002; 2002US-0383830P.
 PR 01-OCT-2002; 2002US-00262839.

XX (CURA-) CURAGEN CORP.

PI Alsbrook JP, Anderson DM, Boldog FL, Burgess CE, Catterton E;
 PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;
 PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;
 PI Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ;
 PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;
 XX WPI; 2003-381625/36.
 DR N-PSDB; ACC72077.
 XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
 PT dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.
 PS Claim 1; Page 107; 487pp; English.

CC The present invention relates to novel human NOV proteins and their
 CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
 CC proteins are useful in manufacturing a medicament for treating a syndrome
 CC associated with a human disease. The NOV proteins and coding sequences
 CC may be used to diagnose, treat or prevent metabolic disorders such as
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune
 CC disorders, haematopoietic disorders and various dyslipidaemias

XX Sequence 1120 AA;

Query Match 100.0%; Score 125; DB 6; Length 1120;
 Best Local Similarity 100.0%; Pred. No. 6.2e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRRSARRRRRPGDPTPKVLE 24
 |||||
 Db 1097 KLGFRRSARRRRRPGDPTPKVLE 1120
 |||||

RESULT 3

AAB30929

ID AAB30929 standard; protein; 1188 AA.

XX AAB30929;

XX 02-APR-2001 (first entry)

XX Amino acid sequence of a human alpha11 integrin chain.

XX Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;
 KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
 KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;
 KW osteoporosis; cartilage damage; bone damage; cartilage.

XX Homo sapiens.

Key	Location/Qualifiers
Peptide	1..22
Region	/note= "signal peptide"
Domain	951..972
	/note= "leucine zipper"
	1142..1164
	/note= "transmembrane domain"

XX WO2000075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-SB001135.

XX 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

DR WPI: 2001-071061/08.
DR N-PSDB; AAC68971.
XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit
PT alpha 11 in association with subunit beta, useful for treating muscle
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.
XX
PS Disclosure; Fig 2a-c; 79pp; English.
XX
CC The present sequence represents a human integrin subunit, designated
CC alpha11. The alpha11 polynucleotide and polypeptide are useful as markers
CC of cell target molecules, such as fibroblasts, muscle cells,
CC chondrocytes, osteoblasts, mesenchymally derived cells or stem cells.
CC They are also used for determining the differential-stage of cells during
CC differentiation, development in pathological conditions, in tissue
CC regeneration, in transplantation or in therapeutic and physiological
CC repair of tissues. The pathological conditions involving subunit alpha11
CC are selected from damage of cells, muscle dystrophy, fibrosis, wound
CC healing, trauma, rheumatoid arthritis, osteoarthritis and osteoporosis,
CC damage of cartilage and bone, and cartilage and bone diseases. The
CC polypeptide is useful for detecting the formation of cartilage during
CC embryonic development, for detecting physiological therapeutic repair of
CC cartilage and muscle, for selection and analysis, or for sorting,
CC isolating or purification of chondrocytes and muscle cells, for detecting
CC regeneration of cartilage or chondrocytes during transplantation of
CC cartilage or chondrocytes during transplantation of cartilage or
CC chondrocytes, respectively, or of muscle or muscle cells during
CC transplantation of muscle or muscle cells, respectively, and for studies
CC of differentiation or chondrocytes or muscle cells
XX
SQ Sequence 1188 AA;

Query Match 100.0%; Score 125; DB 4; Length 1188;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFPSARRRRPGLDTPKVL 24
DB 1165 KLGFPSARRRRPGLDTPKVL 1188

RESULT 4
AAU14467
ID AAU14467 standard; protein; 1188 AA.
AC AAU14467;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human novel protein #338.
XX
KW Human, novel protein; Antianemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
FN WO200155437-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US002623.
XX
PR 25-JAN-2000; 2000US-00491404.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX

DR WPI: 2001-451939/48.
DR N-PSDB; AAS22772.
XX Isolated polypeptides useful for treating anti-inflammatory diseases,
PT nervous system disorders, and for regenerating bone and cartilage.
XX
PS Example 4; Page 828-831; 894pp; English.
XX
CC The invention relates to polynucleotides encoding novel human proteins or
CC their active domains. The polypeptides, polynucleotides and antibodies
CC raised against the polypeptides are used in a method of treatment of a
CC mammal and prevention of disorders caused by the aberrant protein
CC expression or activity. The polypeptides can be used as molecular weight
CC markers, food supplements, and in antibody production. The polypeptides
CC are used to identify compounds which bind to the polypeptides.
CC Polynucleotides of the invention are used as probes and primers, for
CC sequencing, for chromosome or gene mapping, in the production of
CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene
CC therapy. Polypeptides of the invention can be used to target drugs to a
CC tumour, in assays to determine biological activity, to raise
CC antibodies/elicit an immune response, to determine quantitative protein
CC levels, as tissue markers, and to isolate receptors or ligands.
CC Polypeptides of the invention may also be useful in treating platelet
CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
CC ligament and/or nerve tissue, wound healing, treating burns, promoting
CC the proliferation, differentiation and survival of stem cells, as a
CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
CC diseases, nervous system disorders, and infection. The present sequence
CC represents a protein of the invention
XX
SQ Sequence 1188 AA;

Query Match 100.0%; Score 125; DB 4; Length 1188;
Best Local Similarity 100.0%; Pred. No. 6.5e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFPSARRRRPGLDTPKVL 24
DB 1165 KLGFPSARRRRPGLDTPKVL 1188

RESULT 5
AAU14231
ID AAU14231 standard; protein; 1188 AA.
AC AAU14231;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human novel protein #102.
XX
KW Human, novel protein; Antianemic; osteopathic; antiinflammatory;
KW immunomodulatory; cytostatic; neuroprotective; vulnary; nootropic;
KW anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
KW antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;
KW thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
KW Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
KW tissue regeneration; immune disorder.
XX
OS Homo sapiens.
XX
FN WO200155437-A2.
XX
PD 02-AUG-2001.
XX
PF 25-JAN-2001; 2001WO-US002623.
XX
PR 25-JAN-2000; 2000US-00491404.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX

PA (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 FI WPI; 2001-451939/48.
 XX N-PSDB; AAS22536.
 DR
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 PT nervous system disorders, and for regenerating bone and cartilage.
 PT
 XX Example 4; Page 578-581; 894pp; English.
 PS
 XX The invention relates to polynucleotides encoding novel human proteins or
 CC their active domains. The polypeptides, polynucleotides and antibodies
 CC raised against the polypeptides are used in a method of treatment of a
 CC mammal and prevention of disorders caused by the aberrant protein
 CC expression or activity. The polypeptides can be used as molecular weight
 CC markers, food supplements, and in antibody production. The polypeptides
 CC are used to identify compounds which bind to the polypeptides.
 CC Polynucleotides of the invention are used as probes and primers, for
 CC sequencing, for chromosome or gene mapping, in the production of
 CC recombinant proteins, and in generating anti-sense DNA or RNA and in gene
 CC therapy. Polypeptides of the invention can be used to target drugs to a
 CC tumour, in assays to determine biological activity, to raise
 CC antibodies/ elicit an immune response, to determine quantitative protein
 CC levels, as tissue markers, and to isolate receptors or ligands.
 CC Polypeptides of the invention may also be useful in treating platelet
 CC disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 CC ligament and/or nerve tissue, wound healing, treating burns, promoting
 CC the proliferation, differentiation and survival of stem cells, as a
 CC contraceptive, treating osteoporosis and osteoarthritis, anaemia,
 CC Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 CC sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 CC fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
 CC versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
 CC diseases, nervous system disorders, and infection. The present sequence
 CC represents a protein of the invention
 XX
 SQ Sequence 1188 AA;
 Query Match 100.0%; Score 125; DB 4; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KLGFFRSARRRPGDPTPKVLE 24
 Db 1165 KLGFFRSARRRPGDPTPKVLE 1188
 RESULT 6
 AAB50085
 ID AAB50085 standard; protein; 1188 AA.
 AC AAB50085;
 XX
 XX 19-MAR-2001 (first entry)
 DT
 DE Human A259.
 XX
 XX Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
 KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
 KW rheumatoid arthritis.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..1141 /label= Extracellular_domain
 FT Peptide 1..22 /label= Signal_peptide
 FT Protein 22..1188 /label= Mature_protein
 FT Domain 39..174

FT
 FT /label= Integrin_alphasubunit_repeat_domain_#1
 FT 115..157
 FT /label= Integrin_alphasubunit_repeat_domain_#2
 FT 164..345
 FT /label= I_domain
 FT 367..392
 FT /label= Integrin_alphasubunit_repeat_domain_#3
 FT 421..455
 FT /label= Integrin_alphasubunit_repeat_domain_#4
 FT 478..516
 FT /label= Integrin_alphasubunit_repeat_domain_#5
 FT 540..575
 FT /label= Integrin_alphasubunit_repeat_domain_#6
 FT 602..640
 FT /label= Integrin_alphasubunit_repeat_domain_#7
 FT 1142..1164
 FT /label= Transmembrane_domain
 FT 1165..1188
 FT /label= Cytoplasmic_domain
 XX
 PN WO200073339-A1.
 XX
 XX 07-DEC-2000.
 XX
 XX 15-MAY-2000; 2000WO-US013262.
 PF
 XX 28-MAY-1999; 99US-00322790.
 PR
 XX 27-APR-2000; 2000US-00561263.
 PR
 XX (MILL-) MILLENNIUM PHARM INC.
 PA
 PI Pan Y, Lora JM;
 XX
 XX WPI; 2001-041142/05.
 DR
 DR N-PSDB; AAC91901, AAC91902.
 XX
 PT Nucleic acid encoding alpha-integrin subunits, useful for treatment and
 PT diagnosis of fibrosis, e.g. of the liver.
 PS
 XX Claim 8; Fig 1; 164pp; English.
 XX
 CC The present sequence is human integrin alpha subunit, A259. A259 is
 CC homologous with the alpha1 and alpha10 integrin subunits and is
 CC overexpressed in fibrosis. A259 is implicated in regulation of
 CC proliferation, differentiation and/or function of many different cell
 CC types. Inhibitors of A259 activity are useful for the treatment of liver
 CC disease, particularly fibrosis, and also fibrosis in other organs.
 CC (specifically lung and kidney). In addition, A259 can be used for
 CC treatment and prevention of cancer, osteoporosis, acute myeloid
 CC leukaemia, HIV infection, and rheumatoid arthritis
 XX
 SQ Sequence 1188 AA;
 Query Match 100.0%; Score 125; DB 4; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 KLGFFRSARRRPGDPTPKVLE 24
 Db 1165 KLGFFRSARRRPGDPTPKVLE 1188
 RESULT 7
 AAU10551
 ID AAU10551 standard; protein; 1188 AA.
 XX
 XX AAU10551;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 XX Human A259 polypeptide.
 DE
 XX Human; A259; integrin alpha subunit; integrin alpha 10; secreted protein;

XX SQ Sequence 1188 AA;
 Query Match 100.0%; Score 125; DB 7; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 6.5e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRSARRRRPGLDPTPKVLE 24
 |||||
 DB 1165 KLGFRSARRRRPGLDPTPKVLE 1188

RESULT 9
 AAB25582
 ID AAB25582 standard; protein; 1189 AA.
 XX AC AAB25582;
 XX DT 21-NOV-2000 (first entry)
 XX DE ITGA11 protein encoded by human secreted protein gene #7.
 XX KW Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;
 KW antirheumatic; dermatological; antiproliferative; antiarteriosclerotic;
 KW anticancer; vulnary; antiviral; antibacterial; antifungal;
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;
 KW Crohn's disease; nephritis; hyperproliferative disorder;
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;
 KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.
 XX OS Homo sapiens.
 XX FN WO200029435-A1.
 XX PD 25-MAY-2000.
 XX PF 27-OCT-1999; 99WO-US025031.
 XX PR 28-OCT-1998; 98US-0105971P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;
 PI Greene JW;
 XX WPI; 2000-387742/33.
 XX N-PSDB; AAA80612.
 XX PT Isolated nucleic acid molecules encoding human secreted proteins are used
 PT for the prevention, amelioration and treatment of autoimmune,
 PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,
 PT wounds, and infectious diseases.
 XX PS Claim 1; Fig 19A-F; 803pp; English.
 XX CC The present invention relates to 12 secreted human proteins and the
 CC nucleotide sequences encoding them. The polynucleotide sequences given in
 CC AAA80606-A80623 encode the 12 secreted protein sequences given in
 CC AAB25576-B25593. The human secreted proteins have various activities
 CC dependent on the tissues in which they are expressed. Examples of the
 CC activities of the proteins include: immunosuppressant; anti-inflammatory;
 CC antiarthritic; antirheumatic; dermatological; antiproliferative;
 CC antiarteriosclerotic; anticancer; vulnary; antiviral; antibacterial;
 CC and antifungal activity. The proteins, polypeptides, agonists and
 CC antagonists may be used to treat prevent and/or diagnose various disease,
 CC disorders and conditions examples of which include: immune disorders e.g.
 CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple
 CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,
 CC Crohn's disease and nephritis; hyperproliferative disorders such as
 CC paraneoplasias and purpura; cardiovascular disorders e.g. coronary
 CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The
 CC proteins and polynucleotide sequences may also be used in wound healing

CC and the treatment of infectious diseases. The human secreted protein gene
 CC #7 and protein sequences are represented in sequences AAA80612 and
 CC AAB25582. Secreted protein gene #7 is located at position chromosome 15
 CC q22.3-23. Sequences AAA80652-A80661 represent genes which are related to
 CC the secreted protein gene#7
 XX SQ Sequence 1189 AA;
 Query Match 100.0%; Score 125; DB 3; Length 1189;
 Best Local Similarity 100.0%; Pred. No. 6.6e-09;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFRSARRRRPGLDPTPKVLE 24
 |||||
 DB 1166 KLGFRSARRRRPGLDPTPKVLE 1189

RESULT 10
 ABL12949
 ID ABL12949 standard; protein; 1189 AA.
 XX AC ABL12949;
 XX DT 13-FEB-2002 (first entry)
 XX DE Novel human diagnostic protein #12940.
 XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX OS Homo sapiens.
 XX FN WO200175067-A2.
 XX PD 11-OCT-2001.
 XX PF 30-MAR-2001; 2001WO-US008631.
 XX PR 31-MAR-2000; 2000US-00540217.
 XX PR 23-AUG-2000; 2000US-00649167.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Drmanac RT, Liu C, Tang YT;
 XX WPI; 2001-639362/73.
 XX N-PSDB; AAS77136.
 XX PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX PS Claim 20; SEQ ID NO 43308; 103pp; English.
 XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABL00010-ABL30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1189 AA;

Query Match 100.0%; Score 125; DB 4; Length 1189;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREGLDTPPKVLE 24
DB 1166 KLGFFRSARRRREGLDTPPKVLE 1189

RESULT 11
ABR58364
ID ABR58364 standard; protein; 1189 AA.
XX
AC ABR58364;
XX
DT 07-JUL-2003 (first entry)
XX
DE Human NOV2a.
XX
KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;
KW antiparkinsonian; antihypertensive; gene therapy; metabolic disorder;
KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;
KW neurodegenerative disorder; Alzheimer's disease; immune disorder;
KW haematopoietic disorder.
XX
OS Homo sapiens.
XX
PN WO2003029423-A2.
XX
PD 10-APR-2003.
XX
PF 02-OCT-2002; 2002WO-US031358.
PR 02-OCT-2001; 2001US-0326483P.
PR 03-OCT-2001; 2001US-0327342P.
PR 03-OCT-2001; 2001US-0327917P.
PR 09-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328044P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 22-OCT-2001; 2001US-0341658P.
PR 24-OCT-2001; 2001US-0339286P.
PR 24-OCT-2001; 2001US-0343629P.
PR 29-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 12-APR-2002; 2002US-0371972P.
PR 12-APR-2002; 2002US-0371980P.
PR 17-APR-2002; 2002US-0373261P.
PR 19-APR-2002; 2002US-0373805P.
PR 23-APR-2002; 2002US-0374738P.
PR 16-MAY-2002; 2002US-0381101P.
PR 17-MAY-2002; 2002US-0381635P.
PR 29-MAY-2002; 2002US-0383830P.
PR 01-OCT-2002; 2002US-00262839.
XX
FA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;
PI Edinger SA, Ellerman KD, Gerlach VL, Gorman L, Guo X, Ji W;
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;
PI Rothenberg ME, Shinkens RA, Smithson G, Spyttek KA, Taupier RJ;
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;
XX
WPI; 2003-381625/36.

DR N-PSDB; ACC72076.
XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
PT dyslipidaemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Claim 1; Page 105; 487pp; English.
XX
CC The present invention relates to novel human NOV proteins and their
CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
CC proteins are useful in manufacturing a medicament for treating a syndrome
CC associated with a human disease. The NOV proteins and coding sequences
CC may be used to diagnose, treat or prevent metabolic disorders such as
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
CC disorders such as Alzheimer's disease or Parkinson's disease, immune
CC disorders, haematopoietic disorders and various dyslipidaemias
XX
SQ Sequence 1189 AA;

Query Match 100.0%; Score 125; DB 6; Length 1189;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREGLDTPPKVLE 24
DB 1166 KLGFFRSARRRREGLDTPPKVLE 1189

RESULT 12
ADA27054
ID ADA27054 standard; protein; 1189 AA.
XX
AC ADA27054;
XX
DT 20-NOV-2003 (first entry)
XX
DE Human novel secreted protein from cDNA HOHEY69 #1.
XX
KW cytostatic; antiinflammatory; immunomodulator; neuroprotective;
KW hemostatic; gene therapy; cancer; inflammation; immune disorder;
KW neurological disorder; blood clotting disorder; food additive;
KW preservative; human; secreted protein.
XX
OS Homo sapiens.
XX
PN US2003055231-A1.
XX
PD 20-MAR-2003.
XX
PF 29-OCT-2001; 2001US-00984130.
XX
PR 28-OCT-1998; 98US-0105971P.
PR 27-OCT-1999; 99WO-US025031.
PR 19-APR-2000; 2000US-0198407P.
PR 30-OCT-2000; 2000US-0243792P.
PR 18-APR-2001; 2001US-00636353.
XX
NI (NIJJ/) NI J.
PA (YOUN/) YOUNG P E.
PA (KERN/) KENNY J J.
PA (OLSE/) OLSEN H S.
PA (MOOR/) MOORE P A.
PA (WEIY/) WEI Y.
PA (GREE/) GREENE J M.
PA (RUBE/) RUBEN S M.
PA (LIUD/) LIU D.
PA (CROC/) CROCKER P R.
XX
PI Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
PI Ruben SM, Liu D, Crocker PR;
XX
WPI; 2003-567103/53.

DR N-PSDB; ADA27036.

XX New human secreted nucleic acid molecules and polypeptides, useful for

XX preventing, treating, or ameliorating a medical condition, such as

PT cancer, inflammation, immune disorders, neurological and blood clotting

PT disorders.

XX

PS Claim 11; Fig 19; 454pp; English.

XX

CC The invention relates to an isolated nucleic molecule that is at least

CC 95% identical to 18 human cDNA sequences representing 12 novel genes

CC encoding secreted proteins or a polynucleotide fragment of the cDNA

CC sequence contained in American Type Culture Collection (ATCC) deposit No.

CC defined in the specification, its species homologue, a variant or allelic

CC variant of the polynucleotide having a polynucleotide capable of

CC hybridising under conditions the polynucleotide, where the polynucleotide

CC does not hybridise under stringent conditions to a nucleic acid molecule

CC having a nucleotide sequence of only A or T residues. Also included are

CC recombinant vectors, host cells (for producing the polypeptide), the

CC secreted polypeptide (comprising a sequence that is at least 95%

CC identical to a polypeptide fragment, domain, epitope, full-length

CC protein, variant, allelic variant or species homologue), antibodies that

CC specifically bind to the polypeptides, diagnosing, treating, preventing

CC or ameliorating a medical condition by administering the polynucleotide

CC or the polypeptide, the gene corresponding to the cDNA sequence and

CC identifying an activity in a biological assay (by expressing the cDNA

CC sequence in a cell, isolating the supernatant, and detecting an activity

CC in a biological assay and identifying the protein in the supernatant

CC having the activity). The polypeptides, nucleic acids and antibodies are

CC useful for diagnosing a pathological condition or a susceptibility to a

CC pathological condition, for preventing, treating, or ameliorating a

CC medical condition, such as cancer, inflammation and other immune

CC disorders, neurological and blood clotting disorders (many examples are

CC given in the specification). The nucleic acids are also useful for

CC chromosome identification, radiation hybrid mapping or long-range

CC restriction mapping. The polypeptides and antibodies are useful for

CC providing immunological probes for differential identification of the

CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,

CC agonist or antagonist may also be used as a food additive or preservative

CC to increase or decrease storage capabilities, fat content or other

CC nutritional components. The present is a secreted protein of the

CC invention.

XX

Sequence 1189 AA;

Query Match 100.0%; Score 125; DB 6; Length 1189;

Best Local Similarity 100.0%; Pred. No. 6.6e-09;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFSARRRRPGDPTPKVLE 24

Db 1166 KLGFFSARRRRPGDPTPKVLE 1189

RESULT 13

ADBE63570

ID ADE63570 standard; protein; 1189 AA.

XX

AC ADE63570;

XX

DT 29-JAN-2004 (first entry)

DE Human Protein Q9UXX5, SEQ ID NO 9514.

XX

XX Human; pain; neuronal tissue; gene therapy;

XX Spinal segmental nerve injury; chronic constriction injury; CCI;

XX spared nerve injury; SNI; Chung.

XX

OS Homo sapiens.

XX

XX WO2003016475-A2.

XX

PD 27-FEB-2003.

XX

PF 14-AUG-2002; 2002WO-US025765.

XX

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX

XX (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX

XX Woolf C, D'urso D, Befort K, Costigan M;

PI WPI; 2003-269312/26.

XX

XX New composition comprising two or more isolated polypeptides, useful for

PT preparing a medicament for treating pain in an animal.

XX

XX Claim 1; Page; 1017pp; English.

XX

CC The invention discloses a composition comprising two or more isolated rat

CC or human polynucleotides or a polynucleotide which represents a fragment,

CC derivative or allelic variation of the nucleic acid sequence. Also

CC claimed are a vector comprising the novel polynucleotide, a host cell

CC comprising the vector, a method for identifying a nucleotide sequence

CC which is differentially regulated in an animal subjected to pain and a

CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence

CC that is differentially expressed in neuronal tissue of a first animal

CC subjected to pain, a method for identifying a compound which regulates

CC the expression of a polynucleotide sequence which is differentially

CC expressed in an animal subjected to pain, a method for identifying a

CC compound that regulates the activity of one or more of the

CC polynucleotides, a method for producing a pharmaceutical composition, a

CC method for identifying a compound or small molecule that regulates the

CC activity in an animal of one or more of the polypeptides given in the

CC specification, a method for identifying a compound useful in treating

CC pain and a pharmaceutical composition comprising the one or more

CC polypeptides or their antibodies. The polynucleotide or the compound that

CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene

CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene

CC therapy). The sequence presented is a human protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:

CC The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic form directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences.

XX

Sequence 1189 AA;

Query Match 100.0%; Score 125; DB 7; Length 1189;

Best Local Similarity 100.0%; Pred. No. 6.6e-09;

Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFSARRRRPGDPTPKVLE 24

Db 1166 KLGFFSARRRRPGDPTPKVLE 1189

RESULT 14

ADBE6584

ID ADE86584 standard; protein; 1189 AA.

XX

AC ADE86584;

XX

DT 29-JAN-2004 (first entry)

DE Novel human secreted protein #7.

XX

XX human; secreted protein; cancer; liver disorder; hepatitis;

XX neural disorder; Alzheimer's disease.

XX

OS Homo sapiens.

XX

PN US2003129685-A1.
XX 10-JUL-2003.
XX 18-APR-2001; 2001US-00836353.
XX 28-OCT-1998; 98US-0105971P.
PR 27-OCT-1999; 99WO-US025031.
PR 19-APR-2000; 2000US-0198407P.
XX (NIJ/) NI J.
PA (YOUNG/) YOUNG P E.
PA (KENN/) KENNY J J.
PA (OLSEN/) OLSEN H S.
PA (MOORE/) MOORE P A.
PA (WEI/) WEI Y.
PA (GREENE/) GREENE J M.
PA (RUBEN/) RUBEN S M.
XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
PI Ruben SM;
XX WPI; 2004-020335/02.
DR N-PSDB; ADB86566.
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX Claim 11; SEQ ID NO 35; 380pp; English.
XX The invention relates to an isolated nucleic acid sequence, or its
CC allelic variant, a fragment of the cDNA sequence, or its fragment,
CC domain, epitope or species homologue. The nucleic acid is useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition e.g., cancer, liver disorders such as hepatitis or neural
CC disorders such as Alzheimer's disease. The present sequence represents
CC the amino acid sequence of a novel human secreted protein.
XX Sequence 1189 AA;
Query Match 100.0%; Score 125; DB 8; Length 1189;
Best Local Similarity 100.0%; Pred. No. 6.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 KLGFFRSARRRREPGLDTPKVL 24
Db 1166 KLGFFRSARRRREPGLDTPKVL 1189
RESULT 15
AAB50087
ID AAB50087 standard; protein; 1188 AA.
XX AAB50087;
AC AC
XX 19-MAR-2001 (first entry)
DT DT
XX Murine A259.
DE DE
XX Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
KW rheumatoid arthritis.
XX Mus sp.
OS OS
XX Location/Qualifiers
FH Key
FT Domain 1..1141
FT /label= Extracellular_domain
FT Peptide 1..22
FT /label= Signal_peptide
FT Protein 23..1188
FT /label= Mature_protein

FT Domain 39..74
FT /label= Integrin_alphasubunit_repeat_domain_#1
FT Domain 115..157
FT /label= Integrin_alphasubunit_repeat_domain_#2
FT Domain 164..345
FT /label= I domain
FT Domain 367..392
FT /label= Integrin_alphasubunit_repeat_domain_#3
FT Domain 421..455
FT /label= Integrin_alphasubunit_repeat_domain_#4
FT Domain 478..516
FT /label= Integrin_alphasubunit_repeat_domain_#5
FT Domain 540..575
FT /label= Integrin_alphasubunit_repeat_domain_#6
FT Domain 602..640
FT /label= Integrin_alphasubunit_repeat_domain_#7
FT Domain 1142..1164
FT /label= Transmembrane_domain
FT Domain 1165..1188
FT /label= Cytoplasmic_domain
XX WO200073339-A1.
XX 07-DEC-2000.
XX 15-MAY-2000; 2000WO-US013262.
XX 28-MAY-1999; 99US-00322790.
PR 27-APR-2000; 2000US-00561263.
XX (MILL-) MILLENNIUM PHARM INC.
XX Pan Y, Lora JM;
XX WPI; 2001-041142/05
DR N-PSDB; AAC91904; AAC91905.
XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and
PT diagnosis of fibrosis, e.g. of the liver.
XX Claim 8; Fig 5; 164pp; English.
XX The present sequence is murine integrin alpha subunit, A259. A259 is
CC homologous with the alpha and alpha0 integrin subunits and is
CC overexpressed in fibrosis. A259 is implicated in regulation of
CC proliferation, differentiation and/or function of many different cell
CC types. Inhibitors of A259 activity are useful for the treatment of liver
CC disease, particularly fibrosis, and also fibrosis in other organs
CC (specifically lung and kidney). In addition, A259 can be used for
CC treatment and prevention of cancer, osteoporosis, acute myeloid
CC leukaemia, HIV infection, and rheumatoid arthritis
XX Sequence 1188 AA;
Query Match 74.4%; Score 93; DB 4; Length 1188;
Best Local Similarity 70.8%; Pred. No. 0.00022;
Matches 17; Conservative 4; Mismatches 3; Indels 0; Gaps 0;
QY 1 KLGFFRSARRRREPGLDTPKVL 24
Db 1165 KLGFFRSARRRREPGLDTPKVL 1188

Search completed: September 21, 2004, 12:59:13
Job time : 6.57542 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:03:59 ; Search time 4.32402 Seconds
(without alignments)
1782.414 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1188
Perfect score: 125
Sequence: 1 KLGFPSARRRRPGLDTPKVLK 24

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1342398 seqs, 321133274 residues

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA.*
1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep.*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
16: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
17: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
18: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	125	100.0	1120	12	US-10-262-839-6
2	125	100.0	1188	15	US-10-291-265-338
3	125	100.0	1188	15	US-10-291-265-810
4	125	100.0	1189	10	US-09-984-130-35
5	125	100.0	1189	10	US-09-836-353A-35
6	125	100.0	1189	12	US-10-262-839-4
7	92	73.6	545	10	US-09-866-050A-590
8	92	73.6	688	10	US-09-866-050A-624
9	52	41.6	420	16	US-10-437-963-106810
10	51.5	41.2	2614	16	US-10-437-963-188563
11	51	40.8	161	10	US-09-764-891-5091
12	51	40.8	161	14	US-10-205-428-457
13	50	40.0	68	16	US-10-437-963-125024
14	50	40.0	115	12	US-10-425-114-48965
15	50	40.0	117	12	US-10-424-599-170608

Sequence 40558, A
Sequence 180056,
Sequence 220139,
Sequence 185606,
Sequence 127748,
Sequence 2410, Ap
Sequence 2, Appli
Sequence 188616,
Sequence 244523,
Sequence 113972,
Sequence 173508,
Sequence 67432, A
Sequence 26, Appl
Sequence 26, Appl
Sequence 30, Appl
Sequence 22651, A
Sequence 115461,
Sequence 158468,
Sequence 130197,
Sequence 194014,
Sequence 183985,
Sequence 113590,
Sequence 252098,
Sequence 48187, A
Sequence 216264,
Sequence 161117,
Sequence 127750,
Sequence 60791, A
Sequence 49090, A
Sequence 125679,
US-10-425-114-40558
US-10-437-963-180056
US-10-424-599-220139
US-10-437-963-185606
US-10-437-963-127748
US-10-094-749-2410
US-09-834-765-2
US-10-437-963-188616
US-10-424-599-244523
US-10-437-963-113972
US-10-437-963-173508
US-10-425-114-67432
US-09-905-173-26
US-10-060-432-26
US-10-369-493-30
US-10-369-493-22651
US-10-437-963-114561
US-10-437-963-158468
US-10-437-963-130197
US-10-437-963-194014
US-10-437-963-183985
US-10-437-963-113590
US-10-424-599-252098
US-10-767-701-48187
US-10-424-599-216264
US-10-437-963-161117
US-10-437-963-127750
US-10-425-114-60791
US-10-425-114-49090
US-10-437-963-125679

ALIGNMENTS

RESULT 1
US-10-262-839-6
; Sequence 6, Application US/10262839
; Publication No. US20040038877A1
; GENERAL INFORMATION:

; APPLICANT: Alsobrook, John,
; APPLICANT: Anderson, David W.,
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Catterton, Elina,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Leach, Martin,
; APPLICANT: Li, Li,
; APPLICANT: Miller, Charles,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Smithson, Glenda,
; APPLICANT: Spvtek, Kimberly,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Vernhet, Corine,
; APPLICANT: Voss, Edward,
; APPLICANT: Zerhusen, Brian,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462A
; CURRENT APPLICATION NUMBER: US/10/262,839
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: Curaseq1ist version 0.1
; SEQ ID NO 6
; LENGTH: 1120
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-839-6

Query Match 100.0%; Score 125; DB 12; Length 1120;
Best Local Similarity 100.0%; Pred. No. 3.6e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||||
DB 1097 KLGFFRSARRRREPGLDPTPKVLE 1120

RESULT 2

US-10-291-265-338
; Sequence 338, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; CURRENT FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 338
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-338

Query Match 100.0%; Score 125; DB 15; Length 1188;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||||
DB 1165 KLGFFRSARRRREPGLDPTPKVLE 1188

RESULT 3

US-10-291-265-810

; Sequence 810, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; CURRENT FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 810
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-810

Query Match 100.0%; Score 125; DB 15; Length 1188;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||||
DB 1165 KLGFFRSARRRREPGLDPTPKVLE 1188

RESULT 4

US-09-984-130-35
; Sequence 35, Application US/09984130
; Publication No. US20030055231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PP489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-35

Query Match 100.0%; Score 125; DB 10; Length 1189;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||||
DB 1166 KLGFFRSARRRREPGLDPTPKVLE 1189

RESULT 5

US-09-836-353A-35

; Sequence 35, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489PL
; CURRENT APPLICATION NUMBER: US/09/836,353A
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-35

Query Match 100.0%; Score 125; DB 10; Length 1189;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||
Db 1166 KLGFFRSARRRREPGLDPTPKVLE 1189

RESULT 6

US-10-262-839-4

; Sequence 4, Application US/10262839

; Publication No. US20040038877A1

; GENERAL INFORMATION:

; APPLICANT: Alsbrook, John,
; APPLICANT: Anderson, David W.,
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Catterton, Elina,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojie,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Leach, Martin,
; APPLICANT: Li, Li,
; APPLICANT: Miller, Charles,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Smithson, Glenda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Taupier, Raymond, jr.,
; APPLICANT: Vernet, Corine,
; APPLICANT: Voss, Edward,
; APPLICANT: Zerhusen, Brian,
; APPLICANT: Zhong, Mei
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD
; FILE REFERENCE: 21402-462A
; CURRENT APPLICATION NUMBER: US/10/262,839
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/326,483
; PRIOR FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: 60/327,917
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,029
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/381,101
; PRIOR FILING DATE: 2002-05-16
; PRIOR APPLICATION NUMBER: 60/371,972
; PRIOR FILING DATE: 2002-04-12
; PRIOR APPLICATION NUMBER: 60/327,342
; PRIOR FILING DATE: 2001-10-05
; PRIOR APPLICATION NUMBER: 60/328,044
; PRIOR FILING DATE: 2001-10-09
; PRIOR APPLICATION NUMBER: 60/328,849
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/374,738
; PRIOR FILING DATE: 2002-04-23
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 367
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 4
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-262-839-4

Query Match 100.0%; Score 125; DB 12; Length 1189;
Best Local Similarity 100.0%; Pred. No. 3.8e-09;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||
Db 1166 KLGFFRSARRRREPGLDPTPKVLE 1189

RESULT 7

US-09-866-050A-500

; Sequence 500, Application US/09866050A

; Publication No. US20030040471A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D. Isolated From Skin Cells
; TITLE OF INVENTION: Compositions and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-500

Query Match 73.6%; Score 92; DB 10; Length 545;
Best Local Similarity 70.8%; Pred. No. 9.2e-05;
Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

Qy 1 KLGFFRSARRRREPGLDPTPKVLE 24
|||
Db 522 KLGFFKSARKRREPGLGVPELE 545

RESULT 8

US-09-866-050A-624

; Sequence 624, Application US/09866050A

; Publication No. US20030040471A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene

APPLICANT: Murison, James G.
APPLICANT: Kumble, Krishanand D.
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SOFTWARE: Fast-Seq for Windows Version 4.0
SEQ ID NO 624
LENGTH: 688
TYPE: PRT
ORGANISM: Rat
US-09-866-050A-624

Query Match 73.6%; Score 92; DB 10; Length 688;
Best Local Similarity 70.8%; Pred. No. 0.00012;
Matches 17; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 KLGFFSARRRRPGLDPTPKVLE 24
|||:||||:||||:||||:||||:
Db 665 KLGFFSARRRRPGLDPTPKVLE 688

RESULT 9
US-10-437-963-106810
Sequence 106810, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 106810
LENGTH: 420
TYPE: PRT
FEATURE:
ORGANISM: Oryza sativa
OTHER INFORMATION: Clone ID: PAT_MRT4530_11220C.1.pap
US-10-437-963-106810

Query Match 41.6%; Score 52; DB 16; Length 420;
Best Local Similarity 45.0%; Pred. No. 38;
Matches 9; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 2 LGGFFSARRRRPGLDPTPK 21
|||:||||:||||:||||:||||:
Db 73 LGGFFSARRRRPGLDPTPK 92

RESULT 10
US-10-437-963-188563
Sequence 188563, Application US/10437963
Publication No. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping

TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SOFTWARE: Fast-Seq for Windows Version 4.0
SEQ ID NO 188563
LENGTH: 2614
TYPE: PRT
ORGANISM: Oryza sativa
OTHER INFORMATION: Clone ID: PAT_MRT4530_8515SC.1.pap
US-10-437-963-188563

Query Match 41.2%; Score 51.5; DB 16; Length 2614;
Best Local Similarity 46.9%; Pred. No. 2.9e+02;
Matches 15; Conservative 2; Mismatches 6; Indels 9; Gaps 1;

QY 2 LGGFFSARRRRPGLDPTPKVLE 24
|||:||||:||||:||||:||||:
Db 959 LGGFFSARRRRPGLDPTPKVLE 990

RESULT 11
US-09-764-891-5091
Sequence 5091, Application US/09764891
Publication No. US20030077809A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5091
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (107)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (136)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (139)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (143)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (148)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (156)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-5091

Query Match 40.8%; Score 51; DB 10; Length 161;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 6 RSARRRRPGLDPTPKVLE 24
|||:||||:||||:||||:||||:
Db 35 RSARRRRPGLDPTPKVLE 53

RESULT 12
US-10-205-428-457
Sequence 457, Application US/10205428

Publication No. US20030108907A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P117C1
CURRENT APPLICATION NUMBER: US/10/205,428
CURRENT FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: 09/764,892
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1019
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 457
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (107)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (136)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (139)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (143)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (148)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (156)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-205-428-457
Query Match 40.8%; Score 51; DB 14; Length 161;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 6 RSARRRREPGLDPTPKVLE 24
Db 35 RSRRRRQPLDTPSGPPLLE 53
RESULT 13
US-10-437-963-125024
; Sequence 125024, Application US/10437963
; Publication No. US20040123343A1
Publication No. US20030108907A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437,963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 125024
LENGTH: 68
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(68)
OTHER INFORMATION: unsure at all Xaa locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_27708C.1.pep
US-10-437-963-125024
Query Match 40.0%; Score 50; DB 16; Length 68;
Best Local Similarity 66.7%; Pred. No. 12;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 7 SARRRREPGLDPTPK 21
Db 29 SARRRRPPRRPTPR 43
RESULT 14
US-10-425-114-48965
; Sequence 48965, Application US/10425114
; Publication No. US20040034888A1
GENERAL INFORMATION:
APPLICANT: Liu, Jingdong
APPLICANT: Zhou, Yihua
APPLICANT: Kovalic, David K.
APPLICANT: Screen, Steven E.
APPLICANT: Tabaska, Jack E.
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53313)B
CURRENT APPLICATION NUMBER: US/10/425,114
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 73128
SEQ ID NO 48965
LENGTH: 115
TYPE: PRT
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: LIB4570-009-H8_FLI.pep
US-10-425-114-48965
Query Match 40.0%; Score 50; DB 12; Length 115;
Best Local Similarity 61.5%; Pred. No. 20;
Matches 8; Conservative 4; Mismatches 1; Indels 0; Gaps 0;
QY 6 RSARRRREPGLDP 18
Db 72 RAARRRDFGVQP 84
RESULT 15
US-10-424-599-170608
; Sequence 170608, Application US/10424599

Publication No. US20030108907A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: P117C1
CURRENT APPLICATION NUMBER: US/10/205,428
CURRENT FILING DATE: 2002-07-26
PRIOR APPLICATION NUMBER: 09/764,892
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1019
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 457
LENGTH: 161
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc feature
LOCATION: (107)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (136)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (139)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (143)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (148)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
FEATURE:
NAME/KEY: misc feature
LOCATION: (156)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-205-428-457
Query Match 40.8%; Score 51; DB 14; Length 161;
Best Local Similarity 57.9%; Pred. No. 20;
Matches 11; Conservative 2; Mismatches 6; Indels 0; Gaps 0;
QY 6 RSARRRREPGLDPTPKVLE 24
Db 35 RSRRRRQPLDTPSGPPLLE 53
RESULT 13
US-10-437-963-125024
; Sequence 125024, Application US/10437963
; Publication No. US20040123343A1

```

; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170608
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(117)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125074C.1.pap
US-10-424-599-170608

```

```

Query Match      40.0%; Score 50; DB 12; Length 117;
Best Local Similarity 58.8%; Pred. No. 20;
Matches 10; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```

```

Qy      4 FERSARRRREPGLDPTP 20
      |||||
Db      50 FFRKRESHEPELAPPP 66

```

```

Search completed: September 21, 2004, 13:29:02
Job time : 5.32402 secs

```

28 46 36.8 298 4 US-09-252-991A-31102 Sequence 31102, A
29 46 36.8 318 4 US-09-252-991A-22868 Sequence 22868, A
30 46 36.8 329 4 US-09-252-991A-26989 Sequence 26989, A
31 46 36.8 349 4 US-09-252-991A-23455 Sequence 23455, A
32 46 36.8 438 2 US-08-577-492-38 Sequence 38, Appl
33 46 36.8 438 2 US-09-079-630-38 Sequence 38, Appl
34 46 36.8 481 1 US-08-286-856C-2 Sequence 2, Appl
35 46 36.8 481 1 US-08-472-831-2 Sequence 40, Appl
36 46 36.8 501 2 US-08-577-492-40 Sequence 40, Appl
37 46 36.8 501 3 US-09-079-630-40 Sequence 3, Appl
38 46 36.8 506 1 US-08-286-856C-3 Sequence 32, Appl
39 46 36.8 506 2 US-08-472-831-3 Sequence 32, Appl
40 46 36.8 606 2 US-08-577-492-32 Sequence 32, Appl
41 46 36.8 606 3 US-09-079-630-32 Sequence 32, Appl
42 46 36.8 691 4 US-09-252-991A-16809 Sequence 16809, A
43 45.5 36.4 294 4 US-09-252-991A-32345 Sequence 32345, A
44 45 36.0 133 4 US-09-252-991A-17771 Sequence 17771, A
45 45 36.0 148 4 US-09-252-991A-19469 Sequence 19469, A

ALIGNMENTS

RESULT 1
US-09-252-991A-19842
; Sequence 19842, Application US/09252991A
; Patent No. 8551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19842
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19842

Query Match 43.2%; Score 54; DB 4; Length 353;
Best Local Similarity 66.7%; Pred. No. 1.7;
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;
QY 6 RSARRRPPGLDPTP 20
DB 119 RAARRRDPGALPAP 133

RESULT 2
US-09-252-991A-19055
; Sequence 19055, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19055
; LENGTH: 424
; TYPE: PRT

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model
Run on: September 21, 2004, 12:54:33; Search time 1.27374 Seconds
(without alignments)
972.742 Million cell updates/sec

Title: US-09-980-403-2_COPY_1165_1188
Perfect score: 125
Sequence: 1 KLGFPSARRRRPPGLDPTPKVLE 24

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/prodata/2/aaa/5A-COMB.pep:*
2: /cgn2_6/prodata/2/aaa/5B-COMB.pep:*
3: /cgn2_6/prodata/2/aaa/5A-COMB.pep:*
4: /cgn2_6/prodata/2/aaa/5B-COMB.pep:*
5: /cgn2_6/prodata/2/aaa/PCUTUS-COMB.pep:*
6: /cgn2_6/prodata/2/aaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	54	43.2	353	4 US-09-252-991A-19842 Sequence 19842, A
2	52	41.6	424	4 US-09-252-991A-19055 Sequence 19055, A
3	50	41.6	690	4 US-09-252-991A-24903 Sequence 24903, A
4	51	40.8	272	4 US-09-252-991A-31371 Sequence 31371, A
5	50	40.0	425	4 US-08-577-492-38 Sequence 38, Appl
6	49	39.2	236	4 US-09-252-991A-30071 Sequence 30071, A
7	48	38.4	162	4 US-09-252-991A-24838 Sequence 24838, A
8	48	38.4	373	2 US-08-599-171A-26 Sequence 26, Appl
9	48	38.4	373	2 US-08-646-590B-26 Sequence 26, Appl
10	48	38.4	373	3 US-09-069-226-26 Sequence 26, Appl
11	48	38.4	373	3 US-09-412-184-26 Sequence 26, Appl
12	48	38.4	573	4 US-09-252-991A-22802 Sequence 22802, A
13	48	38.4	885	4 US-09-252-991A-228129 Sequence 228129, A
14	47	37.6	88	4 US-09-543-681A-6510 Sequence 6510, Ap
15	47	37.6	110	4 US-09-252-991A-17182 Sequence 17182, A
16	47	37.6	166	4 US-09-252-991A-21902 Sequence 21902, A
17	47	37.6	256	4 US-09-252-991A-23580 Sequence 23580, A
18	47	37.6	310	4 US-09-252-991A-23221 Sequence 23221, A
19	47	37.6	559	4 US-08-252-991A-24592 Sequence 24592, A
20	47	37.6	600	4 US-09-252-991A-24130 Sequence 24130, A
21	46	36.8	212	4 US-09-252-991A-31401 Sequence 31401, A
22	46	36.8	222	1 US-07-688-352C-46 Sequence 46, Appl
23	46	36.8	222	5 PCT-US91-02714-43 Sequence 43, Appl
24	46	36.8	222	5 PCT-US91-02714-54 Sequence 54, Appl
25	46	36.8	223	2 US-08-474-379C-46 Sequence 46, Appl
26	46	36.8	253	3 US-09-146-249A-46 Sequence 46, Appl
27	46	36.8	253	3 US-08-206-189B-46 Sequence 46, Appl

```

; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (176),(216)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-19055

Query Match      41.6%; Score 52; DB 4; Length 424;
Best Local Similarity 66.7%; Pred. No. 4.2;
Matches 10; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  9  RRRREPGLDTPKVL 23
    |||||
Db   56  RRRRRPGDPLPRL 70

RESULT 3
US-09-252-991A-24903
; Sequence 24903, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24903
; LENGTH: 690
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (140),(161)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-252-991A-24903

Query Match      41.6%; Score 52; DB 4; Length 690;
Best Local Similarity 52.8%; Pred. No. 7.1;
Matches 10; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

QY  3  GFFRSARRRRREPGLDTPK 21
    |||||
Db   23  GIRRQRRRRPGLDPE 41

RESULT 4
US-09-252-991A-31371
; Sequence 31371, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 31371
; LENGTH: 272
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31371

```

```

Query Match      40.8%; Score 51; DB 4; Length 272;
Best Local Similarity 64.3%; Pred. No. 3.7;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY  7  SARRRREPGLDPTP 20
    |||||
Db   176  TARRRRPGASPNP 189

RESULT 5
US-09-252-991A-26326
; Sequence 26326, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26326
; LENGTH: 425
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26326

Query Match      40.0%; Score 50; DB 4; Length 425;
Best Local Similarity 61.1%; Pred. No. 8.5;
Matches 11; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY  6  RSARRRRREPGLDTPPKVL 23
    |||||
Db   15  RSAGRRRRSPGLRQDPDL 32

RESULT 6
US-09-252-991A-30071
; Sequence 30071, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30071
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30071

Query Match      39.2%; Score 49; DB 4; Length 296;
Best Local Similarity 64.3%; Pred. No. 8.2;
Matches 9; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY  10  RRRREPGLDTPPKVL 23
    |||||
Db   1  RRRDFDRDPGLL 14

RESULT 7
US-09-252-991A-24838

```


; Sequence 24838, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24838
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24838

Query Match 38.4%; Score 48; DB 4; Length 162;
Best Local Similarity 50.0%; Pred. No. 6.1;
Matches 10; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

QY 1 KLGFFRSARRRREPGLDPTP 20
Db 86 RLAAFRPHARRQPGKGGT 105

RESULT 8
US-09-599-171A-26
; Sequence 26, Application US/08599171A
; Patent No. 5814473
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
; ADDRESSEE: CECCHI, STEWART & OLSTEIN
; STREET: 6 BECKER FARM ROAD
; CITY: ROSELAND
; STATE: NEW JERSEY
; COUNTRY: USA
; ZIP: 07068
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 INCH DISKETTE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WORD PERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,171A
; FILING DATE: Concurrently
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HERRON, CHARLES J.
; REGISTRATION NUMBER: 28,019
; REFERENCE/DOCKET NUMBER: 331400-38
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-994-1700
; TELEFAX: 201-994-1744
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 AMINO ACIDS
; TYPE: AMINO ACID
; TOPOLOGY: LINEAR
; MOLECULE TYPE: PROTEIN
; US-09-599-171A-26

Query Match 38.4%; Score 48; DB 2; Length 373;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 13 EPGLDPTPKVLE 24
Db 32 EPDLSPKVM 43

RESULT 9
US-08-646-590B-26
; Sequence 26, Application US/08646590B
; Patent No. 5962283
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; APPLICANT: Swanson, Ronald V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,590B
; FILING DATE: 08-May-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/599,171
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US97/01094
; FILING DATE: 21-January-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Haile, Ph.D., Lisa A.
; REGISTRATION NUMBER: 38,347
; REFERENCE/DOCKET NUMBER: 09010/017001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/678-5070
; TELEFAX: 619/678-5099
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 373 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-08-646-590B-26

Query Match 38.4%; Score 48; DB 2; Length 373;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 13 EPGLDPTPKVLE 24
Db 32 EPDLSPKVM 43

RESULT 10
US-09-069-226-26
; Sequence 26, Application US/09069226
; Patent No. 6013509
; GENERAL INFORMATION:
; APPLICANT: WARREN, Patrick V.
; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES

```
/
/ NUMBER OF SEQUENCES: 32
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: CARELLA, BYRNE, BAIN, GILFILLAN,
/ ADDRESSEE: CECCHI, STEWART & OLSTEIN
/ STREET: 6 BECKER FARM ROAD
/ CITY: ROSELAND
/ STATE: NEW JERSEY
/ COUNTRY: USA
/ ZIP: 07068
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: 3.5 INCH DISKETTE
/ COMPUTER: IBM PS/2
/ OPERATING SYSTEM: MS-DOS
/ SOFTWARE: WORD PERFECT 5.1
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/069,226
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION NUMBER: 08/599,171
/ FILING DATE:
/ ATTORNEY/AGENT INFORMATION:
/ NAME: HERRON, CHARLES J.
/ REGISTRATION NUMBER: 28,019
/ REFERENCE/DOCKET NUMBER: 331400-38
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 201-994-1700
/ TELEFAX: 201-994-1744
/ INFORMATION FOR SEQ ID NO: 26:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 373 AMINO ACIDS
/ TYPE: AMINO ACID
/ TOPOLOGY: LINEAR
/ MOLECULE TYPE: PROTEIN
/ FRAGMENT TYPE:
US-09-069-226-26

Query Match 38.4%; Score 48; DB 3; Length 373;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 13 EPGLDTPKYLE 24
Db 32 EPDLESPKWE 43

RESULT 11
US-09-412-184-26
/ Sequence 26, Application US/09412184
/ Patent No. 6268188
/ GENERAL INFORMATION:
/ APPLICANT: Warren, Patrick V.
/ TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
/ NUMBER OF SEQUENCES: 42
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Fish & Richardson, P.C.
/ STREET: 4225 Executive Square, Suite 1400
/ CITY: La Jolla
/ STATE: CA
/ COUNTRY: US
/ ZIP: 92037
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: Windows95
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/412,184
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION NUMBER: US/08/646,590
/ FILING DATE: 08-May-1996

Query Match 38.4%; Score 48; DB 3; Length 373;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 8; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 13 EPGLDTPKYLE 24
Db 32 EPDLESPKWE 43

RESULT 12
US-09-252-991A-22802
/ Sequence 22802, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ PRIOR FILING DATE: 1998-07-27
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 22802
/ LENGTH: 573
/ TYPE: PRT
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22802

Query Match 38.4%; Score 48; DB 4; Length 573;
Best Local Similarity 60.0%; Pred. No. 24;
Matches 9; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 6 RSARRRREPGLDPTP 20
Db 358 RSARCQRWPGVEPVP 372

RESULT 13
US-09-252-991A-26129
/ Sequence 26129, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ CURRENT FILING DATE: 1999-02-18
```

Matches 9; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
QY 6 RSARRRRPGLDPTP 20
| | | | | | | | | |
Db 15 RPRRRRRDPGGRPF 29

Search completed: September 21, 2004, 13:06:20
Job time : 2.27374 secs

PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 26129
LENGTH: 885
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26129

Query Match 38.4%; Score 48; DB 4; Length 885;
Best Local Similarity 64.3%; Pred. No. 38;
Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 7 SARRRRPGLDPTP 20
| | | | | | | | | |
Db 476 AGRRTPEGLRDPD 489

RESULT 14
US-09-543-681A-6510
Sequence 6510, Application US/09543681A
Patent No. 6605709
GENERAL INFORMATION:
APPLICANT: GARY BRETON
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 2709.1002-001
CURRENT APPLICATION NUMBER: US/09/543,681A
CURRENT FILING DATE: 2000-04-05
PRIOR APPLICATION NUMBER: US 60/128,706
PRIOR FILING DATE: 1999-04-09
NUMBER OF SEQ ID NOS: 8344
SEQ ID NO 6510
LENGTH: 88
TYPE: PRT
ORGANISM: Proteus mirabilis
US-09-543-681A-6510

Query Match 37.6%; Score 47; DB 4; Length 88;
Best Local Similarity 38.1%; Pred. No. 4.5;
Matches 8; Conservative 7; Mismatches 6; Indels 0; Gaps 0;
QY 3 GFFRSARRRRPGLDPTPKVL 23
| | | | | | | | | |
Db 1 GWFTRSRKQYKSKKKII 21

RESULT 15
US-09-252-991A-17182
Sequence 17182, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 17182
LENGTH: 110
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17182

Query Match 37.6%; Score 47; DB 4; Length 110;
Best Local Similarity 60.0%; Pred. No. 5.7;

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:24:50 ; Search time 27 Seconds
(without alignments)

4232.427 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRGLVVALSLWPGFT.....FRSARRRRRGLDPTPKVLE 1188

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 283366 seqs, 96191526 residues

Word size : 0

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : PIR_78.*

1: pir1.*

2: pir2.*

3: pir3.*

4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	15	1.3	272	2 A5348	integrin alpha-1
2	15	1.3	1151	2 A5226	integrin alpha-1 c
3	15	1.3	1180	2 A35854	integrin alpha-1 c
4	11	0.9	1170	2 I45914	integrin alpha 2 s
5	11	0.9	1181	2 A33998	integrin alpha-2 c
6	9	0.8	74	2 I51524	integrin alpha-2 s
7	9	0.8	315	2 A43567	permease [imported
8	9	0.8	371	1 T13407	3-isopropylmalate
9	9	0.8	607	2 S0658	legumin - Gnetum g
10	9	0.8	1178	2 S44142	VIA-2 protein homo
11	8	0.7	20	2 A0822	cytochrome P450 PB
12	8	0.7	42	2 A34259	cytochrome P450mt4
13	8	0.7	76	2 I51527	integrin alpha 5 s
14	8	0.7	76	2 A43537	heat-stable antigen
15	8	0.7	76	2 I53107	CD24 precursor - r
16	8	0.7	80	2 A48996	B cell surface ant
17	8	0.7	103	2 A08856	conserved hypother
18	8	0.7	127	2 B75301	hypothetical prote
19	8	0.7	141	2 S15785	heat-stable antigen
20	8	0.7	224	2 E71228	hypothetical prote
21	8	0.7	226	2 S76800	hypothetical prote
22	8	0.7	234	1 S15102	eosinophil major b
23	8	0.7	264	2 S22090	catechol O-methyl
24	8	0.7	308	2 A26937	conserved hypother
25	8	0.7	356	2 F97419	BH1459 conserved h
26	8	0.7	360	2 A85016	hypothetical prote
27	8	0.7	367	2 S19172	cytochrome P450 2B
28	8	0.7	387	1 DYH0D4	dopamine receptor
29	8	0.7	402	1 S23860	chloramphenicol re

30	8	0.7	415	2 G83568	probable permease
31	8	0.7	442	2 AG3504	dihydrofolate synt
32	8	0.7	444	2 T01721	hypothetical prote
33	8	0.7	487	2 T47107	benzaldehyde dehyd
34	8	0.7	491	1 O4RTPB	cytochrome P450 2B
35	8	0.7	491	1 O4RTF2	cytochrome P450 2B
36	8	0.7	500	2 B1047	testosterone 16alp
37	8	0.7	547	2 AC1314	malolactic enzyme
38	8	0.7	547	2 AC1686	malolactic enzyme
39	8	0.7	605	1 W1WLEB	El protein - bovin
40	8	0.7	606	2 A72429	oligopeptide ABC t
41	8	0.7	614	2 A69845	Na+/H+ antiporter
42	8	0.7	620	1 W1WLS2	El protein - bovin
43	8	0.7	685	2 AC0527	ferriochrome transp
44	8	0.7	697	2 T16306	hypothetical prote
45	8	0.7	739	2 A88445	protein C26E6.10 l
46	8	0.7	849	2 B83349	probable C1pA/B-ty
47	8	0.7	1041	2 T31437	integrin alpha cha
48	8	0.7	1065	2 T25069	hypothetical prote
49	8	0.7	1146	2 S40311	integrin - fruit f
50	8	0.7	1153	1 RWHUIC	cell surface glyco
51	8	0.7	1163	1 RWHUIC	cell surface glyco
52	8	0.7	1179	2 A53213	integrin alpha-E c
53	8	0.7	1366	2 S57664	IgA-specific metal
54	8	0.7	1460	2 D81675	polymorphic membra
55	8	0.7	1664	2 S67250	DNA-directed RNA p
56	8	0.7	1802	2 H88444	protein C26E6.12 l
57	8	0.7	2238	1 RRVUBY	genome polyprotein
58	7	0.6	47	2 D81686	hypothetical prote
59	7	0.6	74	2 T47376	hypothetical prote
60	7	0.6	89	2 I49515	B144 protein B - m
61	7	0.6	94	2 A99774	probable excisiona
62	7	0.6	99	2 F85636	probable excisiona
63	7	0.6	103	2 G75257	hypothetical prote
64	7	0.6	104	2 PH0141	integrin alpha cha
65	7	0.6	106	4 S57386	hypothetical prote
66	7	0.6	107	2 UC1127	major allergen cha
67	7	0.6	109	2 C56413	major allergen Fel
68	7	0.6	110	1 INRBT1	insulin precursor
69	7	0.6	110	1 IPRT11	insulin 1 precurs
70	7	0.6	110	2 S30268	protein hdeA precu
71	7	0.6	110	2 F91177	protein hdeA precu
72	7	0.6	110	2 G86023	protein hdeA precu
73	7	0.6	114	2 C95870	hypothetical prote
74	7	0.6	115	2 D75434	hypothetical prote
75	7	0.6	132	2 B81215	conserved hypother
76	7	0.6	133	2 H75386	hypothetical prote
77	7	0.6	139	1 C45051	lampxin 2 precurs
78	7	0.6	139	2 A38612	insulin-like growt
79	7	0.6	141	1 HACH1	hemoglobin alpha-D
80	7	0.6	141	2 S56103	hemoglobin alpha-D
81	7	0.6	142	2 JT0573	retinoic acid-indu
82	7	0.6	142	2 G87647	hypothetical prote
83	7	0.6	143	2 T05564	hypothetical prote
84	7	0.6	154	2 S14947	2S albumin - Brazi
85	7	0.6	154	2 F83074	conserved hypother
86	7	0.6	158	2 G82494	conserved hypother
87	7	0.6	160	1 F70351	NADH2 dehydrogenas
88	7	0.6	160	2 S58214	apoptosis inducer
89	7	0.6	161	2 D71903	hit family protein
90	7	0.6	163	2 S66795	probable membrane
91	7	0.6	163	2 B72225	conserved hypother
92	7	0.6	171	1 B64498	conserved hypother
93	7	0.6	176	2 AC0207	probable exported
94	7	0.6	176	2 T38925	cytochrome oxidase
95	7	0.6	176	2 T46709	hypothetical prote
96	7	0.6	176	2 AH1250	hypothetical prote
97	7	0.6	179	2 A64551	hypothetical prote
98	7	0.6	179	2 D71957	hypothetical prote
99	7	0.6	180	2 E70358	HupE hydrogenase r
100	7	0.6	182	2 E82967	hypothetical prote

ALIGNMENTS

```

RESULT 1
A:Species: Gallus gallus (chicken)
C:Date: 10-Feb-1995 #sequence_revision 10-Feb-1995 #text_change 15-Sep-2003
C:Accession: A55348
R:Kern, A.; Briesewitz, R.; Bank, I.; Marcantonio, E.E.
J. Biol. Chem. 269, 22811-22816, 1994
A:Title: The role of the I domain in ligand binding of the human integrin alpha-1beta-1.
A:Reference number: A55348; MUID:94357930; PMID:7521332
A:Accession: A55348
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-272 <KEP>
A:Cross-references: GB:U01114
F:155-230/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match      1.3%; Score 15; DB 2; Length 272;
Best Local Similarity 100.0%; Pred. No. 1.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
Db 57 DIVIVLDGNSIYPW 71

RESULT 2
A:Species: Homo sapiens (man)
C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 15-Sep-2003
C:Accession: A45226
R:Briesewitz, R.; Epstein, M.R.; Marcantonio, E.E.
J. Biol. Chem. 268, 2989-2996, 1993
A:Title: Expression of native and truncated forms of the human integrin alpha 1 subunit.
A:Reference number: A45226; MUID:93155124; PMID:8428973
A:Accession: A45226
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1151 <BRI>
A:Experimental source: hepatoblastoma cell line HepG2
A:Note: sequence extracted from NCBI backbone (NCBI:P124326)
F:142-317/Domain: von Willebrand factor type A repeat homology <VWA1>

Query Match      1.3%; Score 15; DB 2; Length 1151;
Best Local Similarity 100.0%; Pred. No. 6.5e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
Db 144 DIVIVLDGNSIYPW 158

RESULT 3
A:Species: Rattus norvegicus (Norway rat)
C:Date: 23-Oct-1990 #sequence_revision 13-Sep-1991 #text_change 15-Sep-2003
C:Accession: A35854; S11243
R:Ignatius, M.J.; Large, T.H.; Houde, M.; Tawil, J.W.; Barton, A.; Esch, F.; Carbonetto, J.
J. Cell Biol. 111, 709-720, 1990
A:Title: Molecular cloning of the rat integrin alpha-1 subunit: a receptor for laminin A
A:Reference number: A35854; MUID:90338125; PMID:2380249
A:Accession: A35854
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1180 <IGN>
A:Cross-references: GB:X52140; NID:G56493; PIDN:CAA36384.1; PID:G56494
C:Keywords: cell adhesion; cytoskeleton; transmembrane protein
F:170-345/Domain: von Willebrand factor type A repeat homology <VWA2>

```

```

Query Match      1.3%; Score 15; DB 2; Length 1180;
Best Local Similarity 100.0%; Pred. No. 6.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
Db 172 DIVIVLDGNSIYPW 186

RESULT 4
A:Species: Bos primigenius taurus (cattle)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 15-Sep-2003
C:Accession: I45914
R:Kamata, T.; Puzon, W.; Takada, Y.
J. Biol. Chem. 269, 9659-9663, 1994
A:Title: Identification of putative ligand binding sites within the I-domain of integrin alpha 2 subunit - bovine (fragment)
A:Reference number: A54402; MUID:94193647; PMID:7511592
A:Accession: I45914
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1170 <KAM>
A:Cross-references: GB:L25886; NID:G439695; PIDN:AA859255.1; PID:G439696
F:161-336/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match      0.9%; Score 11; DB 2; Length 1170;
Best Local Similarity 100.0%; Pred. No. 0.071;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
Db 495 TDVLLVGAPMY 505

RESULT 5
A:Species: Homo sapiens (man)
C:Date: 30-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 15-Sep-2003
C:Accession: A33998; B56793; A53117
R:Takada, Y.; Hemler, M.E.
J. Cell Biol. 109, 397-407, 1989
A:Title: The primary structure of the VLA-2/collagen receptor alpha(2)-subunit (platelet)
A:Reference number: A33998; MUID:89308879; PMID:2545729
A:Accession: A33998
A:Molecule type: mRNA
A:Residues: 1-1181 <TAK>
A:Cross-references: GB:X17033; NID:G33906; PIDN:CAA34894.1; PID:G33907
A:Note: the authors translated the codon GAT for residue 802 as Gln, GTC for residue 803
R:Camel, B.; Parmentier, S.; Leung, L.L.; McGregor, J.L.
Biochem. J. 279, 419-425, 1991
A:Title: Separation of important new platelet glycoproteins (GPIa, GPIC, GPIC*, GPIIa and GPIb)
A:Reference number: A56793; MUID:92061944; PMID:1953640
A:Accession: B56793
A:Molecule type: protein
A:Residues: 30-43 <CAT>
R:Zutter, M.M.; Santoro, S.A.; Painter, A.S.; Tsung, Y.L.; Gafford, A.
J. Biol. Chem. 269, 463-469, 1994
A:Title: The human alpha-2 integrin gene promoter. Identification of positive and negative regulatory elements
A:Reference number: A53117; MUID:94103255; PMID:8276836
A:Accession: A53117
A:Molecule type: DNA
A:Residues: 1-16, 'V', 18-21 <ZUT>
A:Cross-references: GB:I24121; NID:G400342; PIDN:AAA16619.2; PID:G458353
A:Note: authors translated the codon GTA for residue 17 as Leu
C:Genetics
A:Gene: GDB:ITGA2; CD49B
A:Cross-references: GDB:I28031; OMIM:192974
A:Map position: 5q11.1-5q11.2

```

C;Keywords: cell adhesion; glycoprotein; heterodimer; transmembrane protein
F;1-29/Domain: signal sequence #status predicted <SIG>
F;10-1133/Domain: extracellular #status predicted <EXT>
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 01-Feb-2002
C;Accession: T43407; T39850; T39871
F;173-347/Domain: von Willebrand factor type A repeat homology <VWA2>
F;1134-1154/Domain: transmembrane #status predicted <TM>
F;1155-1181/Domain: intracellular #status predicted <CYT>
F;105,112,343,432,460,475,699,1057,1074,1081/Binding site: carbohydrate (Asn) (covalent)

Query Match 0.9%; Score 11; DB 2; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.072; 0; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 495 TDVLLVGAPMY 505
||| |||||
Db 506 TDVLLVGAPMY 516
||| |||||

RESULT 6
IS1524
Integrin alpha 2 subunit - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 15-Sep-2003
C;Accession: IS1524
R;Whittaker, C.A.; DeSimone, D.W.
Development 117, 1239-1249, 1993
A;Title: Integrin alpha subunit mRNAs are differentially expressed in early Xenopus embryo
A;Reference number: IS1524; MUID:94008528; PMID:8404528
A;Accession: IS1524
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-74 <WHI>
A;Cross-references: GB:L10186; NID:G214538; PIDN:AAA16246.1; PID:G214539

Query Match 0.8%; Score 9; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 0.67; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 495 TDVLLVGAP 503
||| |||||
Db 14 TDVLLVGAP 22
||| |||||

RESULT 7
AH3567
permease [imported] - Brucella melitensis (strain 16M)
C;Species: Brucella melitensis
C;Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
C;Accession: AH3567
R;DelVecchio, V.G.; Kaputal, V.; Redkar, R.J.; Patra, G.; Mujex, C.; Los, T.; Ivanova,
; Nazur, M.; Goltzman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letessier,
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
A;Title: The genome sequence of the facultative intracellular pathogen Brucella melitensis
A;Reference number: AD3252; PMID:11756688
A;Accession: AH3567
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-315 <KUR>
A;Cross-references: GB:AE008918; PIDN:AAL53707.1; PID:G17984630; GSPDB:GNC00191
A;Experimental source: strain 16M
C;Genetics:
A;Gene: BMEI10465
A;Map position: II

Query Match 0.8%; Score 9; DB 2; Length 315;
Best Local Similarity 100.0%; Pred. No. 2.3; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1154 LLLALLVLA 1162
||| |||||
Db 283 LLLALLVLA 291
||| |||||

RESULT 8

T43407
3-isopropylmalate dehydrogenase (EC 1.1.1.85) [validated] - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 21-Jul-2000
C;Accession: T43407; T39850; T39871
R;Kikuchi, Y.; Kikazawa, Y.; Shimatake, G.; Yamamoto, M.
Curr. Genet. 14, 375-379, 1988
A;Title: The primary structure of the leu+ gene of Schizosaccharomyces pombe.
A;Reference number: Z22491; MUID:89106270; PMID:3063400
A;Accession: T43407
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-371 <KIK>
A;Cross-references: EMBL:M36910; NID:G173411; PIDN:AAA35316.1; PID:G173412
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Volckaert, G.
submitted to the EMBL Data Library, July 1998
A;Reference number: Z21885
A;Accession: T39850
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-371 <WOO>
A;Cross-references: EMBL:AL031174; PIDN:CAA20106.1; GSPDB:GN000067
A;Experimental source: strain 972h-; cosmid c1A4
R;Volckaert, G.; Lyne, M.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, February 1998
A;Reference number: Z21887
A;Accession: T39871
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 152-371 <VOL>
A;Cross-references: EMBL:AL021746; PIDN:CAA16840.1; GSPDB:GN000067; SPDB:SPBCL1E8.07c
A;Experimental source: strain 972h-; cosmid c1E8
C;Genetics:
A;Gene: leu1; SPDB:SPBCL1E8.07c; SPBCL1A4.02c
A;Map position: 2
C;Function: <IPD1>
A;Description: catalyzes the oxidation of 3-carboxy-2-hydroxy-4-methylpentanoate to 3-carboxy-2-hydroxy-4-methylpentanoate
A;Pathway: leucine biosynthesis
C;Function: <IPD2>
A;Description: EC 1.1.1.85 [validated, MUID:89106270]
C;Superfamily: 3-isopropylmalate dehydrogenase
C;Keywords: oxidoreductase

Query Match 0.8%; Score 9; DB 1; Length 371;
Best Local Similarity 100.0%; Pred. No. 2.7; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 378 DGVLGAVG 386
||| |||||
Db 69 DGVLGAVG 77
||| |||||

RESULT 9
S60658
legumin - Gnetum gnemon
N;Alternate names: 11S globulin; seed storage protein
C;Species: Gnetum gnemon
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 20-Aug-1999
C;Accession: S60658
R;Braun, H.; Horstmann, C.; Baumlein, H.
submitted to the EMBL Data Library, August 1995
A;Description: Legumins of the Gnetatae: characterization and evolutionary relationship
A;Reference number: S60658
A;Accession: S60658
A;Molecule type: mRNA
A;Residues: 1-607 <BRA>
A;Cross-references: EMBL:Z50779; NID:G949870; PIDN:CAA90642.1; PID:G949871
C;Superfamily: glycinin
C;Keywords: seed; storage protein

Query Match 0.8%; Score 9; DB 2; Length 607;
Best Local Similarity 100.0%; Pred. No. 4.2; 0; Indels 0; Gaps 0;
Matches 9; Conservative 0; Mismatches 0;

QY 1153 LLLALLLV 1161
| | | | |
Db 10 LLLALLLV 18

RESULT 10

844142
VWA-2 protein homolog - mouse
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 15-Sep-2003
C:Accession: S44142
R:Edelman, J.M.; Chan, B.M.; Uniyal, S.; Onodera, H.; Wang, D.Z.; Damjanovich, L.; Latze
submitted to the EMBL Data Library, January 1994
A:Description: The mouse VWA-2 homologue supports collagen and laminin adhesion but not
A:Reference number: S44142
A:Accession: S44142
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1178 <EDE>
A:Cross-references: EMBL:Z29987; NID:g473098; PIDN:CAA82877.1; PID:g473099
P:169-344/Domain: von Willebrand factor type A repeat homology <VWA2>

Query Match 0.8%; Score 9; DB 2; Length 1178;
Best Local Similarity 100.0%; Pred. No. 7.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAP 503
| | | | |
Db 503 TDVLLVGAP 511

RESULT 11

A60822
Cytochrome P450 PB-3a - rat (fragment)
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 11-Jun-1999
C:Accession: A60822; I55191
R:Ameliaz, Z.; Narbonne, J.F.; Wolf, C.R.; Robertson, L.W.; Oesch, F.
Biochem. Pharmacol. 37, 3245-3249, 1988
A:Title: Effect of nutritional imbalances on cytochrome P-450 isozymes in rat liver.
A:Reference number: A60822; MUID:88293549; PMID:3041969
A:Accession: A60822
A:Status: preliminary
A:Molecule type: protein
A:Residues: 1-20 <AME>
R:Hashimoto, T.; Matsumoto, T.; Nishizawa, M.; Kawabata, S.
J. Biochem. 103, 487-492, 1988
A:Title: A mutant rat strain deficient in induction of a phenobarbital-inducible form of
A:Reference number: I55191; MUID:88273074; PMID:2839467
A:Accession: I55191
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-20 <RES>
A:Cross-references: GB:D00250; NID:g20727; PIDN:BAA00181.1; PID:g220728
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: electron transfer; heme; monooxygenase; oxidoreductase; transmembrane protein

Query Match 0.7%; Score 8; DB 2; Length 20;
Best Local Similarity 100.0%; Pred. No. 2.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
| | | | |
Db 6 LLLALLLV 13

RESULT 12

A34259
Cytochrome P450mt4, phenobarbital-inducible, mitochondrial, hepatic - rat (fragment)
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Rattus norvegicus (Norway rat)

C:Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 05-Mar-1999
C:Accession: A34259
R:Shayiq, R.M.; Avadhani, N.G.
Biochemistry 29, 866-873, 1990
A:Title: A phenobarbital-inducible hepatic mitochondrial cytochrome P-450 immunochemical
A:Reference number: A34259; MUID:90254127; PMID:2340279
A:Accession: A34259
A:Molecule type: protein
A:Residues: 1-42 <SHA>
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: electron transfer; heme; liver; mitochondrion; monooxygenase; oxidoreductase

Query Match 0.7%; Score 8; DB 2; Length 42;
Best Local Similarity 100.0%; Pred. No. 4.1;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
| | | | |
Db 6 LLLALLLV 13

RESULT 13

I51527
Integrin alpha 5 subunit - African clawed frog (fragment)
C:Species: Xenopus laevis (African clawed frog)
C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 29-Sep-1999
C:Accession: I51527
R:Whittaker, C.A.; Desimone, D.W.
Development 117, 1239-1249, 1993
A:Title: Integrin alpha subunit mRNAs are differentially expressed in early Xenopus embryo
A:Reference number: I51524; MUID:94008528; PMID:8404528
A:Accession: I51527
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-76 <WHI>
A:Cross-references: GB:I10191; NID:g214544; PIDN:AAA16249.1; PID:g214545
C:Superfamily: integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 6.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 RFGSSIAS 548
| | | | |
Db 63 RFGSSIAS 70

RESULT 14

A43537
Heat-stable antigen M1/69-J11d precursor - mouse
N:Alternate names: CD24 protein; nectadlin
C:Species: Mus musculus (house mouse)
C:Date: 06-Nov-1992 #sequence_revision 06-Nov-1992 #text_change 31-Jan-2000
C:Accession: A43537; I48287; S15784; S15783; S43709; S32240; S33129
R:Kay, R.; Takei, F.; Humphries, R.K.
J. Immunol. 145, 1952-1959, 1990
A:Title: Expression cloning of a cDNA encoding M1/69-J11d heat-stable antigens.
A:Reference number: A43537; MUID:90361906; PMID:2118158
A:Accession: A43537
A:Molecule type: mRNA
A:Residues: 1-76 <KAY>
A:Cross-references: GB:M58661; NID:g198985; PIDN:AAA39481.1; PID:g198986
R:Wenger, R.H.; Rochelle, J.M.; Seldin, M.F.; Kohler, G.; Nielsen, P.J.
J. Biol. Chem. 268, 23345-23352, 1993
A:Title: The heat stable antigen (mouse CD24) gene is differentially regulated but has a
A:Reference number: A48976; MUID:94043127; PMID:8226859
A:Accession: I48287
A:Status: translation not shown; translated from GB/EMBL/DDBJ

A:Residues: 1-76 <KAY>
A:Cross-references: EMBL:X72910; NID:g296466; PIDN:CAA51415.1; PID:g296467
R:Wenger, R.H.; Ayane, M.; Bose, R.; Koehler, G.; Nielsen, P.J.
Eur. J. Immunol. 21, 1019-1046, 1991

A;Title: The genes for a mouse hematopoietic differentiation marker called the heat-stab
A;Reference number: S15783; MUID:91209380; PMID:2019286
A;Accession: S15784
A;Molecule type: DNA
A;Residues: 1-76 <WE3>
A;Cross-references: EMBL:X56469; NID:G51439; PIDN:CAA39841.1; PID:G51440
A;Accession: S15783
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 32-76 <WE2>
A;Cross-references: EMBL:X53825
R;Nielsen, P.J.
submitted to the EMBL Data Library, July 1990
A;Reference number: S19111
A;Accession: S43709
A;Molecule type: mRNA
A;Residues: 1-76 <NIE>
A;Cross-references: EMBL:X53825; NID:G55441; PIDN:CAA37822.1; PID:G55442
C;Genetics:
A;Gene: CD24
A;Introns: 23/3
A;Map position: 10
C;Keywords: blocked carboxyl end; glycoprotein; lipoprotein; membrane protein; phosphati
F;1-26/Domain: signal sequence #status predicted <SIG>
F;27-56/Product: heat-stable antigen M169-Jld #status predicted <MAT>
F;57-76/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F;56/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form)

Query Match 0.7%; Score 8; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

QY 1152 GLLLLALL 1159
|||||
DB 12 GLLLLALL 19

RESULT 15
I53107
CD24 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999
C;Accession: I53107; S25146
R;Shirasawa, T.; Akashi, T.; Sakamoto, K.; Takahashi, H.; Maruyama, N.; Hirokawa, K.
Dev. Dyn. 198, 1-13, 1993
A;Title: Gene expression of CD24 core peptide molecule in developing brain and developin
A;Reference number: I53107; MUID:94122434; PMID:8292828
A;Accession: I53107
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-76 <RES>
A;Cross-references: EMBL:Z11663; NID:G55901; PIDN:CAA7731.1; PID:G55902
C;Keywords: phosphatidylinositol linkage

Query Match 0.7%; Score 8; DB 2; Length 76;
Best Local Similarity 100.0%; Pred. No. 6.9; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

QY 1152 GLLLLALL 1159
|||||
DB 12 GLLLLALL 19

RESULT 16
A48996
B cell surface antigen CD24 precursor - human
N;Alternate names: cluster-w4 antigen; signal transducer CD24
C;Species: Homo sapiens (man)
C;Date: 19-Dec-1993 #sequence_revision 02-Aug-1996 #text_change 21-Jul-2000
C;Accession: I56114; A48996; I54201; I72676
R;Kay, R.; Rosten, P.M.; Humphries, R.K.
J. Immunol. 147, 1412-1416, 1991
A;Title: CD24, a signal transducer modulating B cell activation responses, is a very shd

A;Reference number: I56114; MUID:91332450; PMID:1831224
A;Accession: I56114
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-80 <RES>
A;Cross-references: GB:L33930; NID:G500848; PID:G500849; GB:M58664; NID:G180167; PID:G180167
R;Jackson, D.; Waibel, R.; Weber, E.; Bell, J.; Stahel, R.A.
Cancer Res. 52, 5264-5270, 1992
A;Title: CD24, a signal-transducing molecule expressed on human B cells, is a major sur
A;Reference number: A48996; MUID:93007871; PMID:1327504
A;Accession: A48996
A;Molecule type: mRNA
A;Residues: 1-56, 'V', 58-80 <JAC>
A;Cross-references: GB:X69387; GB:S44888; NID:G396167; PIDN:CAA49195.1; PID:G396168
A;Experimental source: small cell lung carcinoma line SW2
A;Note: sequence extracted from NCBI backbone (NCBI:114635, NCBI:114636)
A;Note: both 57-Val and 57-Ala were found in small cell carcinoma line DC571
R;Hough, M.R.; Rosten, P.M.; Sexton, T.L.; Kay, R.; Humphries, R.K.
Genomics 22, 154-161, 1994
A;Title: Mapping of CD24 and homologous sequences to multiple chromosomal loci.
A;Reference number: I54201; MUID:95048364; PMID:7959762
A;Accession: I54201
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 5-11, 'W', 13-43, 'T', 45-80 <RE2>
A;Cross-references: GB:S75311; NID:G933885; PIDN:AA014170.1; PID:G4261870
C;Genetics:
A;Gene: GDB:CD24
A;Cross-references: GDB:383795; OMIM:600074
A;Map position: 6q21-6q21
C;Keywords: B-cell; blocked carboxyl end; glycoprotein; lipoprotein; phosphatidylinosit
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-67/Product: B cell surface antigen CD24 #status predicted <MAT>
F;68-80/Domain: carboxyl-terminal propeptide #status predicted <PRO>
F;67/Modified site: GPI-anchor ethanolamine amidated carboxyl end (Ser) (in mature form)

Query Match 0.7%; Score 8; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 7.3; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

QY 1152 GLLLLALL 1159
|||||
DB 12 GLLLLALL 19

RESULT 17
AF0856
conserved hypothetical protein STY3056 [imported] - Salmonella enterica subsp. enterica
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AF0856
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher
th, T.; Connor, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica sero
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AF0856
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <PAR>
A;Cross-references: GB:AL513382; PIDN:CAD06037.1; PID:G16504004; GSPDB:GN00176
C;Genetics:
A;Gene: STY3056
C;Superfamily: hypothetical protein HI0673

Query Match 0.7%; Score 8; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 9; Mismatches 0; Indels 0; Gaps 0;
Matches 8; Conservative 0

QY 1153 LLLLLALL 1160

Db 6 LLLLLALLV 13
|||||

RESULT 18

B75301
hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 28-Jul-2000
C:Accession: B75301
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: B75301
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-127 <WHI>
A:Cross-references: GB:AE002054; GB:AE000513; NID:g6460010; PIDN:AAF11763.1; PID:g646001
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR2210
A:Map position: 1
C:Superfamily: Deinococcus radiodurans hypothetical protein DR2210

Query Match 0.7%; Score 8; DB 2; Length 127;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
|||||

Db 7 GLLLLALL 14
|||||

RESULT 19

S15785
heat-stable antigen-related hypothetical protein HSA-C - mouse
C:Species: Mus musculus (house mouse)
C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 05-Nov-1999
C:Accession: S15785
R:Wenger, R.H.; Ayane, M.; Bose, R.; Koehler, G.; Nielsen, P.J.
Eur. J. Immunol. 21, 1039-1046, 1991
A:Title: The genes for a mouse hematopoietic differentiation marker called the heat-stab
A:Reference number: S15783; MUID:91209380; PMID:2019286
A:Accession: S15785
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-141 <HEN>
A:Cross-references: EMBL:X56486; NID:g51441; PIDN:CAA39843.1; PID:g51442
A:Note: the authors translated the codon TTC for residue 87 as Ser and AGA for residue 1
A:Note: the authors did not translate the codon for residue 101
C:Genetics:
A:Introns: #status absent
C:Keywords: glycoprotein; membrane protein; phosphatidylinositol linkage

Query Match 0.7%; Score 8; DB 2; Length 141;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
|||||

Db 12 GLLLLALL 19
|||||

RESULT 20

E71228
hypothetical protein PH0087 - Pyrococcus horikoshii
C:Species: Pyrococcus horikoshii
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 21-Jul-2000
C:Accession: E71228
R:Kawarayashi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Sekim

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Tanaka, Y.; Yamazaki, J.; Kushida, N.; Oguchi.
DNA Res. 5, 55-76, 1998
A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
A:Reference number: A71000; MUID:98344137; PMID:9679194
A:Accession: E71228
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-224 <KAW>
A:Cross-references: GB:AP000001; NID:g3236128; PIDN:BAA29156.1; PID:g3256473
A:Experimental source: strain OT3
A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
C:Genetics:
A:Gene: PH0087

Query Match 0.7%; Score 8; DB 2; Length 224;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 512 RGVVYVE 519
|||||

Db 105 RGVVYVE 112
|||||

RESULT 21

S76800
hypothetical protein - Synechocystis sp. (strain PCC 6803)
C:Species: Synechocystis sp.
A:Variety: PCC 6803
C>Date: 25-Apr-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S76800
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, N.;
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda
DNA Res. 3, 109-136, 1996
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis
s.

A:Reference number: S74322; MUID:97061201; PMID:8905231
A:Accession: S76800
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-226 <KAN>
A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BAA18712.1; PID:g165380
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996
C:Superfamily: conserved hypothetical protein MTH747

Query Match 0.7%; Score 8; DB 2; Length 226;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1073 RLVPNQBI 1080
|||||

Db 183 RLVPNQBI 190
|||||

RESULT 22

S15102
eosinophil major basic protein 2 precursor - guinea pig
C:Species: Cavia porcellus (guinea pig)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Jul-2000
C:Accession: S15102; S18501
R:Aoki, I.; Shindoh, Y.; Nishida, T.; Nakai, S.; Hong, Y.M.; Mio, M.; Saito, T.; Tasaka,
FEBS Lett. 282, 56-60, 1991
A:Title: Comparison of the amino acid and nucleotide sequences between human and two gui
A:Reference number: S15102; MUID:91224343; PMID:2026266
A:Accession: S15102
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-234 <AOK>
A:Cross-references: DDBJ:D00817; NID:g3135095; PIDN:BAA00697.1; PID:g220293

A:Accession: S18501
A:Molecule type: protein
A:Residues: 116, 'X', 118-134, 'X', 136-137, 'X', 139-145; 161-176; 181-200 <AOK2>
C:Superfamily: eosinophil major basic protein precursor; C-type lectin homology
C:Keywords: antibiotic; chondroitin sulfate proteoglycan; cytotoxin; eosinophil; glycom

F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-115/Domain: propeptide #status predicted <PRO>
F;116-234/Product: eosinophil major basic protein 2 #status experimental <MAT>
F;117-232/Domain: C-type lectin homology <LCH>
F;24,25/Binding site: carbohydrate (Ser) (covalent) #status predicted
F;59/Binding site: chondroitin sulfate (Ser) (covalent) #status predicted
F;135-232,209-224/disulfide bonds: #status predicted

Query Match 0.7%; Score 8; DB 1; Length 234;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLALLLV 1160
|||||
Db 4 LLALLLV 11

RESULT 23
S22090
catechol O-methyltransferase (EC 2.1.1.6) - rat
N/Alternate names: catechol-O-methyltransferase
C/Species: Rattus norvegicus (Norway rat)
C/Date: 19-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 21-Jul-2000
C/Accession: S22090; JQ0787; I52356
R;Tenhunen, J.; Salminen, M.; Jalanko, A.; Ukkonen, S.; Ulmanen, I.
submitted to the EMBL Data Library, June 1992
A/Description: Structure of the rat catechol-O-methyltransferase gene: Separate promoter
A/Reference number: S22090
A/Accession: S22090
A/Molecule type: DNA
A/Residues: 1-264 <TEN>
A/Cross-references: EMBL:Z12651; NID:G55889; PIDN:CAA78276.1; PID:G55890
R;Salminen, M.; Lundstroem, K.; Tilmann, C.; Savolainen, R.; Kalkkinen, N.; Ulmanen, I.
Gene 93, 241-247, 1990
A/Title: Molecular cloning and characterization of rat liver catechol-O-methyltransferase
A/Reference number: JQ0787; MUID:91033034; PMID:2227437
A/Accession: JQ0787
A/Molecule type: mRNA
A/Residues: 44-264 <SAL>
A/Cross-references: GB:M60753; NID:G203336; PIDN:AAA40881.1; PID:G203337
A/Experimental source: liver
R;Tenhunen, J.; Ulmanen, I.
Biochem. J. 296, 595-600, 1993
A/Title: Production of rat soluble and membrane-bound catechol O-methyltransferase forms
A/Reference number: I52356; MUID:94107221; PMID:8280056
A/Accession: I52356
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-264 <RES>
A/Cross-references: EMBL:Z12651; NID:G55889; PIDN:CAA78276.1; PID:G55890
C/Comment: This enzyme catalyzes the transfer of a methyl group from S-adenosyl-L-methionine to catechol
A/Gene: COMT
C/Keywords: methyltransferase; S-adenosylmethionine; transmembrane protein

Query Match 0.7%; Score 8; DB 2; Length 264;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLAL 1159
|||||
Db 9 GLLLLAL 16

RESULT 24
AG2637
conserved hypothetical protein Atu0498 [imported] - Agrobacterium tumefaciens (strain C58)
C/Species: Agrobacterium tumefaciens
C/Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 18-Nov-2002
C/Accession: AG2637
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClellan

; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A/Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, J.E.W.
A/Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A/Reference number: AB2577; MUID:21608550; PMID:11743193
A/Accession: AG2637
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-308 <KUR>
A/Cross-references: GB:AE008688; PIDN:AAL41517.1; PID:G17738846; GSPDB:GN00186
A/Experimental source: strain C58 (Dupont)
C/Genetics:
A/Gene: Atu0498
A/Map position: circular chromosome

Query Match 0.7%; Score 8; DB 2; Length 308;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVLA 1162
|||||
Db 275 LLALLVLA 282

RESULT 25
F97419
BH1459 conserved hypothetical protein (AF001512) [imported] - Agrobacterium tumefaciens
C/Species: Agrobacterium tumefaciens
C/Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 18-Nov-2002
C/Accession: F97419
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorllo, B.; Goldman
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A/Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens
A/Reference number: A97359; MUID:21608551; PMID:11743194
A/Accession: F97419
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-356 <KUR>
A/Cross-references: GB:AE007869; PIDN:AAK86311.1; PID:G15155427; GSPDB:GN00169
C/Genetics:
A/Gene: AGR_C 881
A/Map position: circular chromosome

Query Match 0.7%; Score 8; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVLA 1162
|||||
Db 323 LLALLVLA 330

RESULT 26
A85016
hypothetical protein AT4G01220 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C/Accession: A85016
R;anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Harbor Laboratory
A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A/Reference number: A85001; MUID:20083488; PMID:10617198
A/Accession: A85016
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-360 <STO>
A/Cross-references: GB:NC_001268; NID:G7267619; PIDN:CAB80931.1; GSPDB:GN00140
C/Genetics:
A/Gene: AT4G01220
A/Map position: 4

benzaldehyde dehydrogenase (NAD) (EC 1.2.1.28) [validated] - Pseudomonas putida plasmid
 C:Species: Pseudomonas putida
 C>Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 03-Jun-2002
 C:Accession: T47107; S13388
 R:Inoue, J.; Shaw, J.P.; Rekik, M.; Harayama, S.
 J. Bacteriol. 177, 1196-1201, 1995
 A:Title: Overlapping substrate specificity of benzaldehyde dehydrogenase (the xylC gene
 Pseudomonas putida.
 A:Reference number: Z24352; MUID:95173094; PMID:7868591
 A:Accession: T47107
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-487 <INO>
 A:Cross-references: EMBL:U15151; NID:9555994; PIDN:AAA66218.1; PID:9555995
 R:Chalmers, R.M.; Keen, J.N.; Fewson, C.A.
 Biochem. J. 273, 99-107, 1991
 A:Title: Comparison of benzyl alcohol dehydrogenases and benzaldehyde dehydrogenases fr
 uene pathway in Pseudomonas putida.
 A:Reference number: S13298; MUID:91113163; PMID:1989592
 A:Accession: S13388
 A:Molecule type: protein
 A:Residues: 1-44 <CHA>
 A:Experimental source: strain MT53
 C:Genetics:
 A:Gene: xylC
 A:Genome: plasmid
 A>Note: TOL plasmid pWMO
 C:Function:
 A:Description: EC 1.2.1.28 [validated, MUID:95173094]
 C:Superfamily: aldehyde dehydrogenase (NAD+); aldehyde dehydrogenase homology
 C:Keywords: oxidoreductase

Query Match 0.7%; Score 8; DB 2; Length 487;
 Best Local Similarity 100.0%; Pred. No. 35; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;

QY 626 ALGNVAIL 633
 DB 171 ALGNVAIL 178

RESULT 34
 OARTPB
 cytochrome P450 2B1 - rat
 N:Alternate names: cytochrome P450 b; cytochrome P450, phenobarbital-inducible
 N:Contains: unspecific monooxygenase (EC 1.14.14.1)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 18-Aug-1982 #sequence_revision 17-May-1996 #text_change 03-Mar-2000
 C:Accession: A00176; A54251; A22363; A29288; S03854; A92255; I54796
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
 Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982
 A:Title: Primary structure of a cytochrome p-450: coding nucleotide sequence of phenoba
 A:Reference number: A93912; MUID:82222224; PMID:6953431
 A:Accession: A00176
 A:Molecule type: mRNA
 A:Residues: 6-491 <PUJ>
 A:Cross-references: EMBL:J00719; NID:9203752; PIDN:AAA1024.1; PID:9203753
 A>Note: the authors translated the codon GAT for residue 166 as Glu, CTG for residue 29
 R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
 Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982
 A:Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenobar
 A:Reference number: A93925
 A:Contents: annotation
 A>Note: the mistranslations shown in reference A93912 are acknowledged
 R:Roberts, E.S.; Hopkins, N.E.; Zaluzec, E.J.; Gage, D.A.; Alworth, W.L.; Hollenberg, P
 Biochemistry 33, 3766-3771, 1994
 A:Title: Identification of active-site peptides from (3)H-labeled 2-ethynylthiophthalene-
 A:Reference number: A54251; MUID:94190899; PMID:8142377
 A:Accession: A54251
 A:Molecule type: protein
 A:Residues: 290-301, 'X', <ROB>
 R:Suwa, Y.; Mizukami, Y.; Sogawa, K.; Fujii-Kuriyama, Y.
 J. Biol. Chem. 260, 7980-7984, 1985

Query Match 0.7%; Score 8; DB 2; Length 442;
 Best Local Similarity 100.0%; Pred. No. 32; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;

QY 615 NEDGLIDL 622
 DB 241 NEDGLIDL 248

RESULT 32
 T01721
 hypothetical protein A_I0002N01.8 - Arabidopsis thaliana
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C>Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 24-Mar-1999
 C:Accession: T01721
 R:Scheet, P.; Maggi, L.
 submitted to the EMBL Data Library, June 1997
 A:Description: The sequence of A. thaliana I0002N01.
 A:Reference number: Z14407
 A:Accession: T01721
 A>Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-444 <SCH>
 A:Cross-references: EMBL:AF007269; NID:92191126; PID:92191131
 A:Experimental source: cultivar Columbia
 C:Genetics:
 A:Map position: 4
 A:Introns: 68/1; 235/3; 302/3; 374/3
 A>Note: A_I0002N01.8

Query Match 0.7%; Score 8; DB 2; Length 444;
 Best Local Similarity 100.0%; Pred. No. 32; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;

QY 1153 LLLLALL 1160
 DB 115 LLLLALL 122

RESULT 33
 T47107

benzaldehyde dehydrogenase (NAD) (EC 1.2.1.28) [validated] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002
 C:Accession: AG3504
 R:DelVecchio, V.G.; Kapatal, R.J.; Redkar, R.J.; Patra, G.; Mufer, C.; Ios, T.; Ivanova,
 ; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AG3504
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-442 <KOR>
 A:Cross-references: GB:AE008917; PIDN:AAL53202.1; PID:917984076; GSPDB:GNO0190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BM612021
 A:Map position: I
 C:Superfamily: folic polyglutamate synthase
 C:Keywords: ligase

Query Match 0.7%; Score 8; DB 2; Length 415;
 Best Local Similarity 100.0%; Pred. No. 30; Mismatches 0; Indels 0; Gaps 0;
 Matches 8; Conservative 0;

QY 1150 LGGLLLLA 1157
 DB 388 LGGLLLLA 395

RESULT 31
 AG3504
 dihydrofolate synthase (EC 6.3.2.12) [imported] - Brucella melitensis (strain 16M)
 C:Species: Brucella melitensis
 C>Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 19-Apr-2002
 C:Accession: AG3504
 R:DelVecchio, V.G.; Kapatal, R.J.; Redkar, R.J.; Patra, G.; Mufer, C.; Ios, T.; Ivanova,
 ; Mazur, M.; Gotsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AG3504
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-442 <KOR>
 A:Cross-references: GB:AE008917; PIDN:AAL53202.1; PID:917984076; GSPDB:GNO0190
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BM612021
 A:Map position: I
 C:Superfamily: folic polyglutamate synthase
 C:Keywords: ligase

A>Title: Gene structure of a major form of phenobarbital-inducible cytochrome P-450 in rat
A:Reference number: A22363; MUID:85234490; PMID:2989270
A:Accession: A22363
A:Molecule type: DNA
A:Residues: 1-31, 'P', 93-204, 'R', 206-327, 'V', 329-356, 'H', 358-391, 'R', 393-415, 'V', 417-433, 'A', 435-450
A:Cross-references: GB:L00320; NID:G203816; PIDN:AAA41046.1; PID:G203818
A:Note: The authors translated the codon CAG for residue 57 as Gly, CTT for residue 92 as Arg, AAA for residue 236 as Leu, ASC for residue 259 as Asn, GTT for residue 328 as T
R:Angarajan, P.N.; Ravishanker, H.; Padmanaban, G.
Biochem. Biophys. Res. Commun. 144, 258-263, 1987
A>Title: Isolation of a cytochrome P-450e gene variant and characterization of its 5' flanking region
A:Reference number: A29298; MUID:87213174; PMID:3579906
A:Accession: A29298
A>Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-57 <RAN>
R:Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.
Arch. Biochem. Biophys. 270, 23-32, 1989
A>Title: Antibodies targeted against hypervariable and constant regions of cytochromes P-450
A:Reference number: S03854; MUID:89192373; PMID:2539047
A:Accession: S03854
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-18; 146-160, 'E', 162-165; 166, 330-361; 362-380; 402-423 <OES>
R:Botelho, L.H.; Ryan, D.E.; Levin, W.
J. Biol. Chem. 254, 5635-5640, 1979
A>Title: Amino acid compositions and partial amino acid sequences of three highly purified or 3-methylcholanthrene
A:Reference number: A92255; MUID:79194111; PMID:109438
A:Accession: A92255
A:Molecule type: protein
A:Residues: 1-3, 'T', 5-22 <BOT>
R:Fujii-Kuriyama, Y.; Mizukami, Y.; Taniguchi, T.; Muramatsu, M.
Int. Symp. Princess Takamatsu Cancer Res. Fund 12, 31-40, 1982
A>Title: Molecular cloning and coding nucleotide sequence of complementary DNA of cytochrome P-450
A:Reference number: 154796; MUID:93160754; PMID:6300027
A:Accession: 154796
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 6-491 <RES>
A:Cross-references: GB:M37134; NID:G203784; PIDN:AAC42028.1; PID:G203785
C:Genetics:
A:Gene: CYP2B1
A:Introns: 57/3; 112/1; 162/1; 215/3; 274/3; 322/1; 384/3; 432/1
C:Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology
C:Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metalloprotein
F/295-458/Domain: cytochrome P450 homology <P45>
F/302/Active site: Thr #status Predicted
F/436/Binding site: heme iron (Cys) (axial ligand) #status predicted
Query Match 0.7%; Score 8; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1153 LLLALLLV 1160
Db 6 LLLALLLV 13
RESULT 35
O4RTP2
N:Alternate names: cytochrome P450 PB-4; cytochrome P450, phenobarbital-inducible; cytochrome P450 2B2 - rat
N:Contains: oxidoreductase (EC 1.-.-.-)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Dec-1986 #sequence revision 17-May-1996 #text change 01-Dec-2000
C:Accession: A21162; A00177; B00176; B92255; S15589; A21872; A32736; S03855; I59060
R:Mizukami, Y.; Sogawa, K.; Suwa, Y.; Muramatsu, M.; Fujii-Kuriyama, Y.
Proc. Natl. Acad. Sci. U.S.A. 80, 3958-3962, 1983
A>Title: Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver.
A:Reference number: A21162; MUID:83247397; PMID:6306654
A:Accession: A21162

A:Molecule type: DNA
A:Residues: 1-472, 'M', 474-491 <MIZ>
A:Cross-references: EMBL:J00728; NID:G203845; PIDN:AAA41056.1; PID:G203847
A:Note: The authors translated the codon AGT for residue 4 as Thr, and ATG for residue 3 as Arg
R:Prev, A.B.; Waxman, D.J.; Kreibich, G.
J. Biol. Chem. 260, 15253-15265, 1985
A>Title: The structure of phenobarbital-inducible rat liver cytochrome P-450 isoenzyme P-450
A:Reference number: A00177; MUID:86059379; PMID:3877725
A:Accession: A00177
A:Molecule type: protein
A:Residues: 1-291, 'P', 293-320, 'AE', 323-475, 'D', 477-491 <PRE>
R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
Proc. Natl. Acad. Sci. U.S.A. 79, 2793-2797, 1982
A>Title: Primary structure of a cytochrome P-450: coding nucleotide sequence of phenobarbital
A:Reference number: A93912; MUID:82222224; PMID:6953431
A:Accession: B00176
A:Molecule type: mRNA
A:Residues: 6-359, 'S', 361-362, 'V', 364-366, 'V', 368-406, 'S', 408-416, 'N', 418, 'A', 420-477, 'C'
A:Note: nucleotide sequence for residues 1-5 is not given
A:Note: The authors translated the codon GAT for residue 166 as Glu, CTG for residue 292 as Arg
R:Fujii-Kuriyama, Y.; Mizukami, Y.; Kawajiri, K.; Sogawa, K.; Muramatsu, M.
Proc. Natl. Acad. Sci. U.S.A. 79, 5443, 1982
A>Title: Primary structure of a cytochrome P450: coding nucleotide sequence of phenobarbital
A:Reference number: A93925
A:Contents: annotation; revisions
A:Note: the mistranslations in reference A93912 are acknowledged
R:Botelho, L.H.; Ryan, D.E.; Levin, W.
J. Biol. Chem. 254, 5635-5640, 1979
A>Title: Amino acid compositions and partial amino acid sequences of three highly purified or 3-methylcholanthrene
A:Reference number: A92255; MUID:79194111; PMID:109438
A:Accession: B92255
A:Molecule type: protein
A:Residues: 1-3, 'T', 5-22 <BOT>
R:LaCroix, D.; Desrochers, M.; Lambert, M.; Anderson, A.
Gene 86, 201-207, 1990
A>Title: Alternative splicing of mRNA encoding rat liver cytochrome P450e (P450IIB21).
A:Reference number: S15589; MUID:90215299; PMID:2323573
A:Accession: S15589
A:Molecule type: mRNA
A:Residues: 105-113, 'P', 115-274, 'VSPAWRE', 275-321, 'E', 323-491 <LAC>
A:Cross-references: EMBL:M34452; NID:G203679; PIDN:AAA41004.1; PID:G203680
A:Note: translation of the nucleotide sequence is not complete
R:Phillips, I.R.; Shephard, E.A.; Ashworth, A.; Rabin, B.R.
Gene 24, 41-52, 1983
A:Reference number: A21872
A:Accession: A21872
A:Molecule type: mRNA
A:Residues: 168-321, 'E', 323-443, 'K', 445-491 <PHI>
R:Atfalter, M.; Anderson, A.
Biochem. Biophys. Res. Commun. 118, 655-662, 1984
A>Title: Segmental homologies in the coding and 3' non-coding sequences of rat liver cytochrome P-450
A:Reference number: A32736; MUID:84153837; PMID:6322758
A:Accession: A32736
A:Molecule type: mRNA
A:Residues: 385-491 <AFF>
A:Cross-references: GB:K01626; NID:G203782; PIDN:AAA41037.1; PID:G203783
R:Oesch, F.; Waxman, D.J.; Morrissey, J.J.; Honscha, W.; Kissel, W.; Friedberg, T.
Arch. Biochem. Biophys. 270, 23-32, 1989
A>Title: Antibodies targeted against hypervariable and constant regions of cytochromes P-450
A:Reference number: S03854; MUID:89192373; PMID:2539047
A:Accession: S03855
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 329-358, 'AS', 361; 362; 363-380; 402-423 <OES>
R:Archison, M.L.; Adesnik, M.
Proc. Natl. Acad. Sci. U.S.A. 83, 2300-2304, 1986
A>Title: Gene conversion in a cytochrome P-450 gene family.
A:Reference number: I59060; MUID:86205943; PMID:3458196
A:Accession: I59060
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 323-431 <RES>

A;Cross-references: GB:M13234; NID:G203848; PIDN:AAA1057.1; PID:G554434

C;Genetics:

A;Gene: CYP2B2

A;Introns: 384/3

C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C;Keywords: alternative splicing; chromoprotein; electron transfer; endoplasmic reticulum

F;295-458/Domain: cytochrome P450 homology <P45>

F;436/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.7%; Score 8; DB 1; Length 491;

Best Local Similarity 100.0%; Pred. No. 35;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160

|||||

6 LLLLLLV 13

RESULT 36

B31047

Testosterone 16alpha-hydroxylase (EC 1.14.14.-) cytochrome P450 2B10 - mouse

N;Alternate names: cytochrome P450CBP; cytochrome P450pf3/46

C;Species: Mus musculus (house mouse)

C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 04-Mar-2000

C;Accession: B31047; A60559

R;Noshiro, M.; Lakso, M.; Kawajiri, K.; Negishi, M.

Biochemistry 27, 6434-6443, 1988

A;Title: Rip locus: regulation of female-specific isozyme (1-P-450-16-alpha) of testosterone

A;Reference number: A31047; MUID:89118235; PMID:3219345

A;Accession: B31047

A;Molecule type: mRNA

A;Residues: 1-500 <N02>

A;Cross-references: EMBL:M21856; NID:G201968; PIDN:AAA40425.1; PID:G201969

A;Note: the authors translated the codon TTA for residue 55 as Phe, ACA for residue 133

as Val, AAA for residue 236 as Asn, AGG for residue 251 as Lys, AAG for residue 253 as

3 as Asp

A;Note: the authors translated the codon GCA for residue 281 as Thr, TTC for residue 283

8 as Arg, ACA for residue 321 as Ala, GTG for residue 331 as Leu, TCA for residue 349 as

s Ala, and GAC for residue 479 as Gly

A;Note: the sequence nucleotide translation from Fig. 8 is inconsistent with the nucleot

R;Bornheim, L.M.; Correia, M.A.

Mol. Pharmacol. 36, 377-383, 1989

A;Title: Purification and characterization of a mouse liver cytochrome P-450 induced by

A;Reference number: A60559; MUID:89384475; PMID:2779523

A;Accession: A60559

A;Molecule type: protein

A;Residues: 1-15 <B0R>

A;Note: this enzyme was induced by cannabidiol

C;Genetics:

A;Gene: CYP2b-10

C;Superfamily: human cytochrome P450 CYP2D6; cytochrome P450 homology

C;Keywords: chromoprotein; electron transfer; endoplasmic reticulum; heme; iron; metallo

F;295-467/Domain: Cytochrome P450 homology <P45>

F;445/Binding site: heme iron (Cys) (axial ligand) #status predicted

Query Match 0.7%; Score 8; DB 2; Length 500;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160

|||||

6 LLLLLLV 13

RESULT 37

AC1314

malolactic enzyme (malate dehydrogenase) homolog lmo1915 [imported] - Listeria monocytog

C;Species: Listeria monocytogenes

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

C;Accession: AC1314

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.

D.; Jones, L.M.; Karst, U.

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma

ok, C.; Schluerter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AC1314

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-547 <GLA>

A;Cross-references: GB:NC_003210; PIDN:CAC99993.1; PID:G16411368; GSPDB:GN00177

A;Experimental source: strain EGD-e

C;Genetics:

A;Gene: lmo1915

C;Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)

Query Match 0.7%; Score 8; DB 2; Length 547;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 MIVITDGE 273

|||||

141 MIVITDGE 148

RESULT 38

AC1686

malolactic enzyme (malate dehydrogenase) homolog lin2029 [imported] - Listeria innocua

C;Species: Listeria innocua

C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 14-Dec-2001

R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloecke

; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H

Science 294, 849-852, 2001

A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; M

ok, C.; Schluerter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland

A;Title: Comparative genomics of Listeria species.

A;Reference number: AB1077; MUID:21537279; PMID:11679669

A;Accession: AC1686

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-547 <GLA>

A;Cross-references: GB:AL592022; PIDN:CAC97259.1; PID:G16414530; GSPDB:GN00178

A;Experimental source: strain Clp11262

C;Genetics:

A;Gene: lin2029

C;Superfamily: malate dehydrogenase (oxaloacetate-decarboxylating)

Query Match 0.7%; Score 8; DB 2; Length 547;

Best Local Similarity 100.0%; Pred. No. 39;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 266 MIVITDGE 273

|||||

141 MIVITDGE 148

RESULT 39

W1WLEB

E1 protein - bovine papillomavirus type 1

C;Species: bovine papillomavirus type 1

C;Date: 17-Dec-1982 #sequence_revision 17-Dec-1982 #text_change 28-Jul-2000

C;Accession: A03663

R;Chen, E.Y.; Howley, P.M.; Levinson, A.D.; Seeburg, P.H.

Nature 299, 529-534, 1982

A;Title: The primary structure and genetic organization of the bovine papillomavirus ty

A;Reference number: A93289; MUID:83012974; PMID:6289124

A;Accession: A03663

A;Molecule type: DNA

A;Residues: 1-605 <CHE>

A;Cross-references: GB:X02346; GB:J02044; GB:M24622; GB:X00473; NID:G60965; PIDN:CAB465

C;Superfamily: papillomavirus E1 protein

C;Keywords: early protein

Query Match
Best Local Similarity 0.7%; Score 8; DB 1; Length 605;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934
DB 337 LAAGSDSN 344

RESULT 40
A72429
oligopeptide ABC transporter, periplasmic oligopeptide-binding protein - Thermotoga maritima
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: A72429
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hickey, Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.; C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome sequencing
A:Reference number: A72200; PMID:99287316; PMID:10360571
A:Accession: A72429
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-606 <ARN>
A:Cross-references: GB:AE001690; GB:AE000512; NID:G4980496; PIDN:AA035125.1; PID:G498051
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0031
C:Superfamily: dipeptide transport protein

Query Match
Best Local Similarity 0.7%; Score 8; DB 2; Length 606;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 9 LLLALLVL 16

RESULT 41
A69845
Na+/H+ antiporter homolog yjBQ - Bacillus subtilis
C:Species: Bacillus subtilis
C:Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 15-Oct-1999
C:Accession: A69845
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bester, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chodura, E.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.; Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Leber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, K.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sironi, A.; Smit, P.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Trosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, K.; Yamano, K.; Yasunoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; PMID:98044033; PMID:9384377
A:Accession: A69845
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-614 <XUN>
A:Cross-references: GB:Z99110; GB:AL009126; NID:G2633472; PIDN:CB013021.1; PID:el183184;
A:Experimental source: strain 168
C:Genetics:
A:Gene: yjBQ

Query Match
0.7%; Score 8; DB 2; Length 614;

Best Local Similarity 100.0%; Pred. No. 43;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1160
DB 313 LLLALLVL 320

RESULT 42
WLWLS2
E1 protein - bovine papillomavirus type 2
C:Species: bovine papillomavirus type 2
C:Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 12-Jun-1998
C:Accession: C31169
R:Groff, D.E.; Mitra, R.; Lancaster, W.D.
Submitted to GenBank, May 1988
A:Reference number: A94519
A:Accession: C31169
A:Molecule type: DNA
A:Residues: 1-620 <GRO>
A:Cross-references: GB:M20219; GB:M19551; NID:G332996
C:Superfamily: papillomavirus E1 protein
C:Keywords: early protein; glycoprotein
F172109.173/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match
Best Local Similarity 0.7%; Score 8; DB 1; Length 620;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934
DB 352 LAAGSDSN 359

RESULT 43
AC0527
ferrichrome transport protein PhbB precursor STY0221 [imported] - Salmonella enterica subsp. enterica serovar typhi
C:Species: Salmonella enterica subsp. enterica serovar typhi
A:Note: this species has also been called Salmonella typhi
C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C:Accession: AC0527
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moul, S.; O'Garra, P.
Nature 413, 848-852, 2001
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhimurium DT104
A:Reference number: AB0502; PMID:11677608
A:Accession: AC0527
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-685 <PAR>
A:Cross-references: GB:AL513382; PIDN:CAD01355.1; PID:gl6501482; GSPDB:GN00176
C:Genetics:
A:Gene: STY0221
C:Superfamily: vitamin B12 transport protein bruc

Query Match
Best Local Similarity 100.0%; Pred. No. 47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1160
DB 36 LLLALLVL 43

RESULT 44
T16306
hypothetical protein F40F4.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T16306
R:Wilson, R.

submitted to the EMBL Data Library, November 1995
A:Description: The sequence of C. elegans cosmid F40F4.
A:Reference number: Z18493
A:Accession: T16306
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-697 <WIL>
A:Cross-references: EMBL:U40420; NID:g1065513; PID:g1065515; PIDN:AAAG1431.1; CESP:F40F4
C:Genetics:
A:Introns: 14/2; 86/3; 135/3; 223/1; 244/1; 331/1; 365/1; 396/1; 435/1; 559/3; 627/2; 69

Query Match 0.7%; Score 8; DB 2; Length 697;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 557 YNDVVVGA 564
DB 593 YNDVVVGA 600

RESULT 45
A88445
protein C26E6.10 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: A88445
R:anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MUID:9089613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C.elegans/ and www.sanger.ac.uk/Projects/C.ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: A88445
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-739 <STC>
A:Cross-references: GB:chr_III; PIDN:AAA21163.1; PID:g532810; GSPDB:GN00021; CESP:C26E6.
C:Genetics:
A:Gene: C26E6.10
A:Map position: 3

Query Match 0.7%; Score 8; DB 2; Length 739;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 430 TSVVSSRQ 437
DB 234 TSVVSSRQ 241

RESULT 46
E83349
probable C1pA/B-type proteinase PA2371 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: E83349
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Micooguchi, S.D.; Warrenner, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A:Reference number: A82950; MUID:2043737; PMID:10984043
A:Accession: E83349
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-849 <STO>
A:Cross-references: GB:AE004663; GB:AE004091; NID:G9948405; PIDN:AAG05759.1; GSPDB:GN001
C:Genetics:
A:Experimental source: strain PA01
A:Gene: PA2371
C:Superfamily: endopeptidase Clp ATP-binding chain

Query Match 0.7%; Score 8; DB 2; Length 849;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 111 GLLLLALL 118

RESULT 47
TJ1437
integrin alpha chain SU2 - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 21-Jan-2000
C:Accession: T31437
R:Hertzler, P.L.; McClay, D.R.
submitted to the EMBL Data Library, May 1998
A:Description: Alpha SU2, a sea urchin integrin which binds laminin.
A:Reference number: Z21035
A:Accession: T31437
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1041 <HER>
A:Cross-references: EMBL:AF067658; NID:g3220240; PID:g3220241; PIDN:AAC3572.1
A:Experimental source: developmental stage embryo
C:Function:
A:Description: binds laminin
C:Superfamily: integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 1041;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 498 LLVGAPMY 505
DB 308 LLVGAPMY 315

RESULT 48
T25068
hypothetical protein T21C9.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T25068
R:McMurray, A.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19977
A:Accession: T25068
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-1065 <WIL>
A:Cross-references: EMBL:Z73098; PIDN:CAA97331.1; GSPDB:GN00023; CESP:T21C9.2
A:Experimental source: clone T21C9
C:Genetics:
A:Gene: CESP:T21C9.2
A:Map position: 5
A:Introns: 67/2; 147/2; 207/3; 242/3; 293/3; 825/2; 897/1; 1032/2

Query Match 0.7%; Score 8; DB 2; Length 1065;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 866 FASLIQKE 873
DB 457 FASLIQKE 464

RESULT 49
S40311
integrin - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 21-Jul-2000
C:Accession: S40311

R;Wehrli, M.; DiAntonio, A.; Fearnley, I.M.; Smith, R.J.; Wilcox, M. Mech. Dev. 43, 21-36, 1993

A:Title: Cloning and characterization of alpha(P51), a novel Drosophila melanogaster integrin alpha chain; neutrophil adherence receptor alpha chain

A:Reference number: S40311; MUID:94059764; PMID:8240969

A:Accession: S40311

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-1146 <MEH>

A:Cross-references: ENBL:X73975; NID:9440143; PIDN:CAAS2155.1; PID:9440144

A:Gene: FlyBase: new

A:Cross-references: FlyBase:FBgn0004456

C:Superfamily: integrin alpha-2b chain

Query Match 0.7%; Score 8; DB 2; Length 1146;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 0; Gaps 0; Indels 0

QY 1151 GGLLLLLL 1158
|||||

Db 1094 GGLLLLLL 1101
|||||

RESULT 50

RWHUB

cell surface glycoprotein CD11b precursor [validated] - human

N:Alternate names: complement receptor type 3 alpha chain; leukocyte adhesion protein Mac-1; integrin alpha chain; neutrophil adherence receptor alpha chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 31-Dec-2000

C:Accession: A31108; A28915; A41600; A30892; A32218; A46526; A26091; 152567

R:Corbi, A.L.; Kishimoto, T.K.; Miller, L.J.; Springer, T.A. J. Biol. Chem. 263, 12403-12411, 1988

A:Title: The human leukocyte adhesion glycoprotein Mac-1 (complement receptor type 3, CD11b)

A:Reference number: A31108; MUID:68315033; PMID:2457584

A:Accession: A31108

A:Molecule type: mRNA

A:Residues: 1-1153 <COR>

A:Cross-references: GB:J03925; NID:G187284; PIDN:AAA59544.1; PID:9307148

A>Note: part of this sequence was confirmed by protein sequencing

R:Arnaout, M.A.; Gupta, S.K.; Pierce, M.W.; Tenen, D.G. J. Cell Biol. 106, 2153-2158, 1988

A:Title: Amino acid sequence of the alpha subunit of human leukocyte adhesion receptor Mac-1

A:Reference number: A28915; MUID:88257215; PMID:2454931

A:Accession: A28915

A:Molecule type: mRNA

A:Residues: 1-499, 501-965, P', 967-1153 <ARN>

A:Cross-references: GB:M18044; GB:J03270; GB:M19664; GB:X07421; NID:G186935; PIDN:AAA594

A>Note: the authors translated the codon TAC for residue 1129 as Thr

A:Title: part of this sequence, including the amino end of the mature protein, was confirmed by protein sequencing

R:Shelley, C.S.; Arnaout, M.A. Proc. Natl. Acad. Sci. U.S.A. 88, 10525-10529, 1991

A:Title: The promoter of the CD11b gene directs myeloid-specific and developmentally regulated expression of the alpha subunit of human leukocyte adhesion receptor

A:Reference number: A41600; MUID:92073318; PMID:1683702

A:Accession: A41600

A:Molecule type: DNA

A:Residues: 1-9 <SHE>

A:Cross-references: GB:M76724; NID:G180018; PIDN:AAA58410.1; PID:9553215

R:Arnaout, M.A.; Remold-O'Donnell, E.; Pierce, M.W.; Harris, P.; Tenen, D.G. Proc. Natl. Acad. Sci. U.S.A. 85, 2776-2780, 1988

A:Title: Molecular cloning of the alpha subunit of human and guinea pig leukocyte adhesion receptor

A:Reference number: A94193; MUID:88130151; PMID:2833753

A:Accession: A30892

A:Molecule type: mRNA

A:Residues: 917-1042 <AR2>

A:Cross-references: GB:M18044

R:Rickstein, D.D.; Hickey, M.J.; Ozols, J.; Baker, D.M.; Back, A.L.; Roth, G.J. Proc. Natl. Acad. Sci. U.S.A. 86, 257-261, 1989

A:Title: cDNA sequence for the alpha subunit of the human neutrophil adherence receptor

A:Reference number: A32218; MUID:89098893; PMID:2563162

A:Accession: A32218

A:Molecule type: mRNA

A:Residues: 9-1153 <HIC>

A:Cross-references: GB:J04145; NID:G189068; PIDN:AAA59903.1; PID:9386975

A>Note: part of this sequence was confirmed by protein sequencing

R:Fleming, J.C.; Pahl, H.L.; Gonzalez, D.A.; Smith, T.F.; Tenen, D.G. J. Immunol. 150, 480-490, 1993

A:Title: Structural analysis of the CD11b gene and phylogenetic analysis of the alpha chain during evolution

A:Reference number: A46526; MUID:93123748; PMID:8419480

A:Accession: A46526

A>Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-499, 501-1153 <FLE>

A:Cross-references: GB:S52227; NID:G263047; PIDN:AA24821.1; PID:G263049

A>Note: the last three bases of intron 13, CAG, are included in some but not all mature mRNAs

R:Pierce, M.W.; Remold-O'Donnell, E.; Todd III, R.F.; Arnaout, M.A. Biochim. Biophys. Acta 874, 368-371, 1986

A:Title: N-terminal sequence of human leukocyte glycoprotein Mol: conservation across species

A:Reference number: A30664; MUID:87076671; PMID:3539202

A:Accession: A30691

A:Molecule type: protein

A:Residues: 17-31 <PIE>

A:Experimental source: granulocytes

R:Pahl, H.L.; Rosmarin, A.G.; Tenen, D.G. Blood 79, 865-870, 1992

A:Title: Characterization of the myeloid-specific CD11b promoter

A:Reference number: 152567; MUID:92144986; PMID:1346576

A:Accession: 152567

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-9 <RES>

A:Cross-references: GB:M84477; NID:G180184; PIDN:AAA51960.1; PID:9553219

C:Comment: A common beta chain (CD18) forms a heterodimer with this chain to form Mac-1

C:Genetics:

A:Gene: GDB:ITGAM; CR3A

A:Cross-references: GDB:120599; OMIM:120980

A:Map position: 16p11.2-16p11.2

A>Note: promoter contains a GATA motif and two Sp1 consensus binding sites

C:Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat homology

C:Keywords: alternative splicing; calcium; cell adhesion; glycoprotein; heterodimer; mag

F:1-16/Domain: signal sequence #status predicted <SIG>

F:17-1153/Product: cell surface glycoprotein CD11b #status experimental <VAT>

F:17-1108/Domain: extracellular #status predicted <EXT>

F:148-318/Domain: von Willebrand factor type A repeat homology <VWA2>

F:465-473/Region: calcium/magnesium binding #status predicted

F:530-538/Region: calcium/magnesium binding #status predicted

F:593-601/Region: calcium/magnesium binding #status predicted

F:1109-1134/Domain: transmembrane #status predicted <TM>

F:1135-1153/Domain: intracellular #status predicted <INT>

F:86,240,391,469,693,697,735,802,881,901,912,941,947,979,994,1022,1045,1051,1076/Binding

Query Match 0.7%; Score 8; DB 1; Length 1153;
Best Local Similarity 100.0%; Pred. No. 74;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLLLL 1158
|||||

Db 1116 GGLLLLLL 1123
|||||

RESULT 51

RWHUB

cell surface glycoprotein CD11c precursor - human

N:Alternate names: leukocyte adhesion receptor p150,95 alpha chain

C:Species: Homo sapiens (man)

C>Date: 31-Dec-1992 #sequence revision 31-Dec-1992 #text change 22-Jun-1999

C:Accession: A36584; A35543; S00864

R:Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A. J. Biol. Chem. 265, 12750-12751, 1990

A:Reference number: A36584

A:Contents: erratum

A:Accession: A36584

A:Molecule type: DNA

A;Residues: 1-1163 <COR>
A;Note: this revision to the sequence from reference A35543 includes the carboxyl end
R;Corbi, A.L.; Garcia-Aguilar, J.; Springer, T.A.
J. Biol. Chem. 265, 2782-2788, 1990
A;Title: Genomic structure of an integrin alpha subunit, the leukocyte p150,95 molecule.
A;Reference number: A35543; MUID:90153906; PMID:2303426
A;Accession: A35543
A;Molecule type: DNA
A;Residues: 1-834 <CO2>
A;Note: this sequence has been revised in reference A36594
R;Corbi, A.L.; Miller, L.J.; O'Connor, K.; Larson, R.S.; Springer, T.A.
EMBO J. 6, 4023-4028, 1987
A;Title: cDNA cloning and complete primary structure of the alpha subunit of a leukocyte
A;Reference number: S00864; MUID:88166645; PMID:3327687
A;Accession: S00864
A;Molecule type: mRNA
A;Residues: 1-755 'L', 757-1163 <CO3>
A;Cross-references: GB:M81695; EMBL:Y00093; NID:G487829; PIDN:AAA59180.1; PID:G487830
A;Note: part of this sequence was confirmed by protein sequencing
C;Comment: A common beta chain (CD18) forms a heterodimer with CD11b to form Mac-1 on my
C;Genetics:
A;Gene: GDB:ITGAX; CD11C
A;Cross-references: GDB:119758; OMIM:151510
A;Map position: 16p11.2-16p11.2
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
C;Keywords: calcium; cell adhesion; glycoprotein; heterodimer; magnesium; tandem repeat;
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1163/Product: cell surface glycoprotein CD11c #status predicted <EXT>
F;20-1107/Domain: extracellular #status predicted <EXT>
F;149-319/Domain: von Willebrand factor type A repeat homology <VWA4>
F;1108-1133/Domain: transmembrane #status predicted <TM>
F;1134-1163/Domain: intracellular #status predicted <INT>
F;61,89,392,697,735,899,939,1050/Binding site: carboxydrate (Asn) (covalent) #status pre

Query Match 0.7%; Score 8; DB 1; Length 1163;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGULLLAL 1158
|||||
Db 1115 GGULLLAL 1122
|||||

RESULT 52
A53213
Integrin alpha-E chain - human
C;Species: Homo sapiens (man)
C;Date: 19-Oct-1995 #sequence_revision 31-Mar-2001 #text_change 31-Mar-2001
C;Accession: A53213
R;Shaw, S.K.; Cepek, K.L.; Murphy, E.A.; Russell, G.J.; Brenner, M.B.; Parker, C.M.
J. Biol. Chem. 269, 6016-6025, 1994
A;Title: Molecular cloning of the human mucosal lymphocyte integrin alpha (E) subunit. Un
A;Reference number: A53213; MUID:94164962; PMID:8119947
A;Accession: A53213
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-1179 <SHA>
A;Cross-references: GB:125851; NID:G457244; PID:G457245
C;Genetics:
A;Gene: GDB:ITGAE
A;Cross-references: GDB:330801
A;Map position: 17p13
C;Superfamily: cell surface glycoprotein CD11b; von Willebrand factor type A repeat hom
F;199-371/Domain: von Willebrand factor type A repeat homology <VWA3>

Query Match 0.7%; Score 8; DB 2; Length 1179;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VLLGAVGA 387
|||||
Db 416 VLLGAVGA 423
|||||

RESULT 53
S57664
IGA-specific metalloendopeptidase homolog sepA precursor - Shigella flexneri
C;Species: Shigella flexneri
C;Date: 19-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 08-Dec-2000
C;Accession: S57664; S69769; S69768
R;Benjelloun-Touimi, Z.; Sansonetti, P.J.; Parsot, C.
submitted to the EMBL Data Library, February 1995
A;Description: Characterization of SepA, the major extracellular protein of Shigella fl
A;Reference number: S57664
A;Accession: S57664
A;Molecule type: DNA
A;Residues: 1-1366 <BEN>
A;Cross-references: EMBL:Z48219; NID:G886952; PIDN:CAA88252.1; PID:G886953
R;Benjelloun-Touimi, Z.; Sansonetti, P.J.; Parsot, C.
Mol. Microbiol. 17, 123-135, 1995
A;Title: SepA, the major extracellular protein of Shigella flexneri: autonomous secreti
A;Reference number: S69768; MUID:96020667; PMID:7476198
A;Accession: S69769
A;Status: nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-66;1077-1366 <BEW>
A;Cross-references: EMBL:Z48219; NID:G886952; PIDN:CAA88252.1; PID:G886953
A;Experimental source: strain N90T (serotype 5)
A;Accession: S69768
A;Molecule type: protein
A;Residues: 57-61; Q', 63-72; 'X', 544-557; 'X', 1057-1068 <BEF>
A;Note: 6-Glu, 14-Thr, 18-Gly and 26-Gln were also found
C;Genetics:
A;Gene: sepA
C;Superfamily: IGA-specific metalloendopeptidase
C;Keywords: extracellular protein
F;1-56/Domain: signal sequence #status predicted <SIG>
F;57-1366/Product: IGA-specific metalloendopeptidase homolog sepA #status predicted <MA

Query Match 0.7%; Score 8; DB 2; Length 1366;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 100 KDNMRLGL 107
|||||
Db 1335 KDNMRLGL 1342
|||||

RESULT 54
DB1675
polymorphic membrane protein B/C family TC0695 [imported] - Chlamydia muridarum (strain
C;Species: Chlamydia muridarum, Chlamydia trachomatis MoPn
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000
C;Accession: DB1675
R;Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.P.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg
Nucleic Acids Res. 28, 1397-1406, 2000
A;Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.
A;Reference number: AB1500; MUID:20150255; PMID:10684935
A;Accession: DB1675
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1460 <TET>
A;Cross-references: GB:AE002338; GB:AE002160; NID:G7190724; PIDN:AAF39511.1; PID:G71907
A;Experimental source: strain N-gg (MoPn)
C;Genetics:
A;Gene: TC0695

Query Match 0.7%; Score 8; DB 2; Length 1460;
Best Local Similarity 100.0%; Pred. No. 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 583 SILKTPKQ 590
|||||
Db 1311 SILKTPKQ 1318
|||||

RESULT 55

S67250
DNA-directed RNA polymerase (EC 2.7.7.6) I 190K chain - yeast (*Saccharomyces cerevisiae*)
N:Alternate names: DNA-directed RNA polymerase A 190K chain; protein O6276; protein YOR3
C:Species: *Saccharomyces cerevisiae*
C>Date: 12-Jul-1996 #sequence revision 12-Jul-1996 #text_change 20-Jun-2000
C:Accession: S67250; A29926; S67396

R/Goffeau, A.; Purnelle, B.
submitted to the Protein Sequence Database, July 1996
A:Reference number: S67246
A:Accession: S67250
A:Molecule type: DNA

A:Residues: 1-1664 <GOF>
A:Cross-references: EMBL:Z75249; NID:gl1420740; PIDN:CAA99665.1; PID:gl1420741; MIPS:YOR34
A:Experimental source: strain S288C
R:Memet, S.; Gouy, M.; March, C.; Sentenac, A.; Buhler, J.M.

J. Biol. Chem. 263, 2830-2839, 1988
A:Title: RPA190, the gene coding for the largest subunit of yeast RNA polymerase A.

A:Reference number: A29926; MUID:88139333; PMID:2830265
A:Accession: A29926

A:Molecule type: DNA

A:Residues: 1-157; T, 159-1664 <MEM>
A:Cross-references: EMBL:J03530; NID:gl172214; PIDN:AAA34890.1; PID:gl172215

R:Purnelle, B.; Goffeau, A.
submitted to the EMBL Data Library, February 1996
A:Description: Nucleotide sequence analysis of a 40 kb segment on the right arm of yeast

ome I genes.

A:Reference number: S67392
A:Accession: S67396

A:Molecule type: DNA

A:Residues: 1-1664 <PUR>

A:Cross-references: EMBL:X95720; NID:gl199839; PIDN:CAA65029.1; PID:gl199844

C:Genetics:

A:Gene: SGD:RPA190

A:Cross-references: SGD:S0005868; MIPS:YOR341w

A:Map position: 15R

C:Superfamily: Trypanosoma DNA-directed RNA polymerase I largest chain

C:Keywords: DNA binding; nucleotidyltransferase; transcription; zinc finger

F:62-78/Region: zinc finger CCHH motif

Query Match 0.7%; Score 8; DB 2; Length 1664;
Best Local Similarity 100.0%; Pred. No. 1e+02; Indels 0; Gaps 0;
Matches 8; Conservative 0; Mismatches 0;

Qy 481 GSEITSVD 488
Db 8 GSEITSVD 15

RESULT 56

H88444
protein C26E6.12 [imported] - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C:Accession: H88444

R:anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating biolog

A:Reference number: A75000; MUID:99069613; PMID:9851916

A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/projects/C_ele

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A:Accession: H88444

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1802 <STO>

A:Cross-references: GB:chr_III; PIDN:AAA21166.1; PID:G532813; GSPDB:GN00021; CESP:C26E6.

C:Genetics:

A:Gene: C26E6.12

A:Map position: 3

Query Match 0.7%; Score 8; DB 2; Length 1802;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 430 TSVSSSRQ 437

Db 1297 TSVSSSRQ 1304

RESULT 57

RVUBY

genome polyprotein - bunyamwera virus

N:Alternate names: L protein

N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)

C:Species: bunyamwera virus

C>Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 11-Jun-1999

C:Accession: A33744

R:Elliot, R.M.

Virology 173, 426-436, 1989

A:Title: Nucleotide sequence analysis of the large (L) genomic RNA segment of Bunyamwera

A:Reference number: A33744; MUID:90085791; PMID:2596023

A:Accession: A33744

A:Molecule type: genomic RNA

A:Residues: 1-2238 <ELL>

A:Cross-references: GB:X14383; NID:G58712; PIDN:CAA32553.1; PID:G58713

C:Genetics:

A:Map position: segment L

C:Superfamily: bunyavirus RNA-directed RNA polymerase

C:Keywords: nucleotidyltransferase

Query Match 0.7%; Score 8; DB 1; Length 2238;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 184 FLINILKK 191

Db 1333 FLINILKK 1340

RESULT 58

DS1666

hypothetical protein TC0766 [imported] - *Chlamydia muridarum* (strain Nigg)

C:Species: *Chlamydia muridarum*, *Chlamydia trachomatis* MoPn

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 11-May-2000

C:Accession: DS1666

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,

C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: DS1666

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-47 <TET>

A:Cross-references: GB:AE002345; GB:AE002160; NID:G7190791; PIDN:AAF39569.1; PID:G719079

A:Experimental source: strain Nigg (MoPn)

C:Genetics:

A:Gene: TC0766

Query Match 0.6%; Score 7; DB 2; Length 47;
Best Local Similarity 100.0%; Pred. No. 46;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLIALL 1159

Db 4 LLLIALL 10

RESULT 59

T47376

hypothetical protein T5C2.40 - *Arabidopsis thaliana*

C:Species: *Arabidopsis thaliana* (mouse-ear cross)

C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 20-Apr-2000

C:Accession: T47376

R:Obermaier, B.; Ottenwaelder, B.; Duchemin, D.; Zeitler, K.; Mewes, H.W.; Lemcke, K.; M

submitted to the Protein Sequence Database, March 2000

A;Reference number: Z24463
A;Accession: T47376
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-74 <OBE>
A;Cross-references: EMBL:AL138664
A;Experimental source: cultivar Columbia; BAC clone T5C2
C;Genetics:
A;Map position: 3
A;Introns: 36/3; 55/3
A;Note: T5C2.40

Query Match 0.6%; Score 7; DB 2; Length 74;
Best Local Similarity 100.0%; Pred. No. 69;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 786 HCVFPLV 792
|||
DB 13 HCVFPLV 19

RESULT 60

I49515
B144 protein B - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999
C;Accession: I49515
R;Tsuge, I.; Shen, F.
Immunogenetics 26, 378-380, 1987
A;Title: A gene in the H-2S; H-2D interval of the major histocompatibility complex which
A;Reference number: I49514; MUID:88031493; PMID:3117682
A;Accession: I49515
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-89 <RES>
A;Cross-references: GB:M18187; NID:g192097; PIDN:AAA37273.1; PID:g192099

Query Match 0.6%; Score 7; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGJLLL 1156
|||
DB 36 LGGJLLL 42

RESULT 61

A99774
probable excisionase [imported] - Escherichia coli (strain O157:H7, substrain RMD 05099)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 18-Jul-2001
C;Accession: A99774
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: A99774
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-94 <RAY>
A;Cross-references: PIDN:BA000007; PIDN:BA034584.1; PID:g13360621; GSPDB:GN00154
A;Experimental source: strain O157:H7, substrain RMD 0509952
C;Genetics:
A;Gene: ECs1161

Query Match 0.6%; Score 7; DB 2; Length 94;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 637 PVQINA 643
|||

DB 18 PVQINA 24

RESULT 62

F85636
probable excisionase for bacteriophage BP-933W xisW [imported] - Escherichia coli (stra.
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C;Accession: F85636
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhe
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85636
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-99 <STO>
A;Cross-references: GB:AE005174; NID:g12514277; PIDN:AAG55554.1; GSPDB:GN00145; UWGP:Z1
A;Experimental source: strain O157:H7, substrain EDL933
C;Genetics:
A;Gene: xisW

Query Match 0.6%; Score 7; DB 2; Length 99;
Best Local Similarity 100.0%; Pred. No. 89;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 637 PVQINA 643
|||
DB 23 PVQINA 29

RESULT 63

G75257
Hypothetical protein - Deinococcus radiodurans (strain R1)
C;Species: Deinococcus radiodurans
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C;Accession: G75257
R;White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; M.
Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A;Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A;Reference number: A75250; MUID:20036896; PMID:10567266
A;Accession: G75257
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <WHI>
A;Cross-references: GB:AE002086; GB:AE000513; NID:g6460395; PIDN:AAF2117.1; PID:g64604
A;Experimental source: strain R1
C;Genetics:
A;Gene: DR2573
A;Map position: 1

Query Match 0.6%; Score 7; DB 2; Length 103;
Best Local Similarity 100.0%; Pred. No. 92;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1157 ALLVLAL 1163
|||
DB 15 ALLVLAL 21

RESULT 64

PH0141
integrin alpha chain - chicken (fragments)
C;Species: Gallus gallus (chicken)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 15-Sep-2003
C;Accession: PH0141
R;Syfrig, J.; Mann, K.; Paulsson, M.
Exp. Cell Res. 194, 165-173, 1991
A;Title: An abundant chick gizzard integrin is the avian alpha1-beta1 integrin heterodi
A;Reference number: PH0141; MUID:91224161; PMID:1851093

```

A:Accession: PH0141
A:Molecule type: protein
A:Residues: 1-104 <SVF>
A:Experimental source: gizzard
C:Comment: Integrins are a family of cellular adhesion receptors that mediate interaction
C:Keywords: cell adhesion

Query Match      0.6%; Score 7; DB 2; Length 104;
Best Local Similarity 100.0%; Pred. No. 93;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      475 QIGSYFG 481
DB      45 QIGSYFG 51
|||||

RESULT 65
S57386
hypothetical protein orf 00954 - Yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 28-Oct-1995 #sequence_revision 05-Sep-1996 #text_change 20-Oct-2000
C:Accession: S57386
R:Zumstein, E.; Pearson, B.M.; Kalogeropoulos, A.; Schweizer, M.
A:Title: A 29.425 kb segment on the left arm of yeast chromosome XV contains more than
A:Reference number: S57374; MUID:96021609; PMID:8533473
A:Accession: S57386
A>Status: nucleic acid sequence not shown; conceptual translation of pseudogene
A:Molecule type: DNA
A:Residues: 1-106 <ZUM>
A:Cross-references: EMBL:X83121; NID:G600461; PIDN:CAA58196.1; PID:G600475
C:Comment: There is no evidence that this sequence is expressed.
C:Genetics:
A:Map position: 15L
C:Keywords: pseudogene

Query Match      0.6%; Score 7; DB 4; Length 106;
Best Local Similarity 100.0%; Pred. No. 94;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      117 SFLACSP 123
DB      17 SFLACSP 23
|||||

RESULT 66
JC1127
major allergen chain 2 precursor, short form - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C:Accession: JC1127
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L.
A:Title: Expression and genomic structure of the genes encoding FdI, the major allergen
A:Reference number: JC1126; MUID:92241678; PMID:1572548
A:Accession: JC1127
A:Molecule type: DNA
A:Residues: 1-107 <GRI>
A:Cross-references: GB:X62478; NID:G395406; PIDN:CAA44345.1; PID:G395407
A:Experimental source: skin
C:Genetics:
A:Gene: Ch2
A:Introns: 21/1; 81/3
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-107/Product: major allergen chain 2, short form #status predicted <MAT>
F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      0.6%; Score 7; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 95;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1157 ALLVLAL 1163
|||||

```

```

DB      4 ALLVLAL 10
|||||

RESULT 67
CS6413
major allergen Fel dI chain 2 precursor - cat
C:Species: Felis silvestris catus (domestic cat)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 01-Dec-2000
C:Accession: CS6413; JC1145
R:Morgenstern, J.P.; Griffith, I.J.; Brauer, A.W.; Rogers, B.L.; Bond, J.F.; Chapman, M.
A:Title: Amino acid sequence of Fel dI, the major allergen of the domestic cat: protein
A:Reference number: A56413; MUID:92052157; PMID:1946388
A:Accession: CS6413
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-109 <MOR>
A:Cross-references: GB:M77341; NID:G163822; PIDN:AA41616.1; PID:G163823
R:Griffith, I.J.; Craig, S.; Pollock, J.; Yu, X.B.; Morgenstern, J.P.; Rogers, B.L.
A:Title: Expression and genomic structure of the genes encoding FdI, the major allergen
A:Reference number: JC1126; MUID:92241678; PMID:1572548
A:Accession: JC1145
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 18-109 <GRI>
A:Experimental source: salivary gland
C:Keywords: glycoprotein
F:1-17/Domain: signal sequence #status predicted <SIG>
F:50/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      0.6%; Score 7; DB 2; Length 109;
Best Local Similarity 100.0%; Pred. No. 97;
Matches      7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1157 ALLVLAL 1163
DB      4 ALLVLAL 10
|||||

RESULT 68
INRB
insulin precursor - rabbit
N:Alternate names: preproinsulin
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 24-Apr-1984 #sequence_revision 23-Aug-1997 #text_change 18-Jun-1999
C:Accession: A53438; A01581
R:Devaskar, S.U.; Giddings, S.J.; Rajakumar, P.A.; Carnaghi, L.R.; Menon, R.K.; Zahn, D.
J. Biol. Chem. 269, 8445-8454, 1994
A:Title: Insulin gene expression and insulin synthesis in mammalian neuronal cells.
A:Reference number: A53438; MUID:94179230; PMID:8132571
A:Accession: A53438
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-110 <DEV>
A:Cross-references: GB:U03610; NID:G467970; PIDN:AAA19033.1; PID:G467971
R:Smith, L.F.
A:Title: Species variation in the amino acid sequence of insulin.
A:Reference number: A90029; MUID:66160119; PMID:5949593
A:Accession: A01581
A:Molecule type: protein
A:Residues: 25-54;90-110 <SMI>
C:Superfamily: insulin
C:Keywords: hormone; pancreas
F:1-24/Domain: signal sequence #status predicted <SIG>
F:25-54/Domain: insulin chain B #status experimental <BCH>
F:25-54;90-110/Product: insulin #status experimental <MAT>
F:57-87/Domain: connecting C peptide #status predicted <CPFP>
F:90-110/Domain: insulin chain A #status experimental <ACH>
F:31-96,43-109,95-100/Disulfide bonds: #status predicted

```

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVL 1161
|||
Db 10 LLALLVL 16

RESULT 69
IPRT1
insulin 1 precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 23-Oct-1981 #sequence_revision 23-Oct-1981 #text_change 24-Sep-1999
C;Accession: A90788; A90789; A94311; B92120; I51945; A01589
R;Cordall, B.; Bell, G.; Tischer, E.; Denoco, F.M.; Ullrich, A.; Pictet, R.; Rutter, W.J.
Cell 18, 533-543, 1979
A;Title: Isolation and characterization of a cloned rat insulin gene.
A;Reference number: A90788; MUID:80045034; PMID:498283
A;Accession: A90788
A;Molecule type: DNA
A;Residues: 1-110 <COR>
A;Cross-references: GB:J00747; NID:G204956; PIDN:AAA41442.1; PID:G204957
R;Lomedico, P.; Rosenthal, N.; Efstratiadis, A.; Gilbert, W.; Kolodner, R.; Tizard, R.
Cell 18, 545-558, 1979
A;Title: The structure and evolution of the two nonallelic rat preproinsulin genes.
A;Reference number: A90789; MUID:80045035; PMID:498284
A;Accession: A90789
A;Molecule type: DNA
A;Residues: 1-110 <JOM>
A;Cross-references: GB:J00747; NID:G204956; PIDN:AAA41442.1; PID:G204957
R;Steiner, D.F.; Clark, J.L.; Nolan, C.; Rubenstein, A.H.; Margoliash, E.; Aten, B.; Oye
Recent Prog. Horm. Res. 25, 207-282, 1969
A;Title: Proinsulin and the biosynthesis of insulin.
A;Reference number: A94311; MUID:70067613; PMID:4311938
A;Accession: A94311
A;Molecule type: protein
A;Residues: 25-54; 90-110 <STE>
R;Tagger, H.S.; Steiner, D.F.
J. Biol. Chem. 247, 7936-7940, 1972
A;Title: Primary structures of the proinsulin connecting peptides of the rat and horse.
A;Reference number: A92120; MUID:73061498; PMID:4640931
A;Accession: B92120
A;Molecule type: protein
A;Residues: 57-87 <TAG>
R;Lomedico, P.T.; Rosenthal, N.; Kolodner, R.; Efstratiadis, A.; Gilbert, W.
Ann. N. Y. Acad. Sci. 343, 425-432, 1980
A;Title: The structure of rat preproinsulin genes.
A;Reference number: I51945; MUID:80240379; PMID:6249167
A;Accession: I51945
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-110 <RES>
A;Cross-references: GB:M25584; NID:G204947; PIDN:AAA41439.1; PID:G204948
C;Genetics:
A;Gene: INS1
C;Superfamily: insulin
C;Keywords: hormone; pancreas
F;1-24/Domain: signal sequence #status predicted <SIG>
F;25-54/Domain: insulin chain B #status experimental <BCH>
F;25-54_90-110/Product: insulin #status experimental <MAT>
F;57-87/Domain: connecting peptide #status experimental <CPEP>
F;90-110/Domain: insulin chain A #status experimental <ACH>
F;31-96,43-109,95-100/Disulfide bonds: #status experimental

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVL 1161
|||
Db 10 LLALLVL 16

RESULT 70

S30268
protein hdeA precursor - Escherichia coli (strain K-12)
N;Alternate names: 10K S protein precursor; hypothetical protein A
C;Species: Escherichia coli
C;Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 01-Mar-2002
C;Accession: S30268; S30266; S47730; A65149
R;Yoshida, T.; Ueguchi, C.; Yamada, H.; Mizuno, T.
Mol. Gen. Genet. 237, 113-122, 1993
A;Title: Function of the Escherichia coli nucleoid protein, H-NS: molecular analysis of
A;Reference number: S30261; MUID:93204884; PMID:8455549
A;Accession: S30268
A;Molecule type: DNA
A;Residues: 1-110 <YOS>
A;Cross-references: EMBL:D11109; NID:G216429; PIDN:BA01883.1; PID:G216431
A;Accession: S30266
A;Molecule type: protein
A;Residues: 22-40 <YOW>
R;Plunkett, G.
submitted to the EMBL Data Library, March 1994
A;Reference number: S47666
A;Accession: S47730
A;Molecule type: DNA
A;Residues: 1-110 <PLU>
A;Cross-references: EMBL:U00039; NID:G466582; PIDN:AAB18486.1; PID:G466647
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.
A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: A65149
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-110 <BLAT>
A;Cross-references: GB:AE000427; GB:U00096; NID:G1789919; PIDN:AAC76535.1; PID:G1789926
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:
A;Gene: hdeA
C;Superfamily: Escherichia coli protein hdeA
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-110/Product: 10K-S protein #status experimental <MAT>

Query Match 0.6%; Score 7; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156
|||
Db 9 LGGLLLL 15

RESULT 71
F91177
protein hdeA precursor - Escherichia coli (strain O157:H7, substrain RIMD 0509952)
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 27-Nov-2001
C;Accession: F91177
R;Hayaishi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91177
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <HAY>
A;Cross-references: GB:BA000007; PIDN:BA837813.1; PID:G13363864; GSPDB:GNO0154
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs4390
C;Superfamily: Escherichia coli protein hdeA

Query Match 0.6%; Score 7; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156
 |||||
 Db 9 LGGLLLL 15

RESULT 72
 G6023
 protein hdeA precursor - Escherichia coli (strain O157:H7, substrain EDL933)
 C:Species: Escherichia coli
 C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 27-Nov-2001
 C:Accession: G6023
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 illet, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A>Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A:Reference number: A85480; MUID:21074935; PMID:11206551
 A:Accession: G6023
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-110 <STO>
 A:Cross-references: GB:AB005174; NID:G12518219; PIDN:AAG58651.1; GSPDB:GN00145; UWGP:249
 A:Experimental source: strain O157:H7, substrain EDL933
 C:Genetics:
 A:Gene: hdeA
 C:Superfamily: Escherichia coli protein hdeA

Query Match 0.6%; Score 7; DB 2; Length 110;
 Best Local Similarity 100.0%; Pred. No. 97;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLL 1156
 |||||
 Db 9 LGGLLLL 15

RESULT 73
 C95870
 hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSymB
 C:Species: Sinorhizobium meliloti
 C>Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 30-Sep-2001
 C:Accession: C95870
 R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernar
 Proc. Natl. Acad. Sci. U.S.A. 98, 9899-9894, 2001
 A>Title: The complete sequence of the 1.683-kb pSymB megaplasmid from the N2-fixing endo
 A:Reference number: A95842; MUID:2136506; PMID:11481431
 A:Accession: C95870
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-114 <KUR>
 A:Cross-references: GB:AL591985; PIDN:CAC48627.1; PID:G15140099; GSPDB:GN00167
 A:Experimental source: strain 1021, megaplasmid pSymB
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pella, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 283, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaud, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A>Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 A:Contents: annotation
 C:Genetics:
 A:Gene: Smb20236
 A:Genome: plasmid

Query Match 0.6%; Score 7; DB 2; Length 114;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 104 RLGLSLA 110

Db 88 RLGLSLA 94
 |||||

RESULT 74
 D75434
 hypothetical protein - Deinococcus radiodurans (strain R1)
 C:Species: Deinococcus radiodurans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: D75434
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
 M.; Shen, M.; Vanathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A>Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: D75434
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-115 <WHI>
 A:Cross-references: GB:AB001962; GB:AE000513; NID:G6458855; PIDN:AAF10702.1; PID:G645886
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR1125
 A:Map position: 1

Query Match 0.6%; Score 7; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159
 |||||
 Db 10 LLLALL 16

RESULT 75
 B81215
 conserved hypothetical inner membrane protein NMA2196 [imported] - Neisseria meningitidis
 C:Species: Neisseria meningitidis
 C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 02-Feb-2001
 C:Accession: B81215; G81792
 R:Retelink, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
 Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
 xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Massignani, V.; Pizza, M.
 Science 287, 1809-1815, 2000
 A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
 A>Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
 A:Reference number: A81000; MUID:20175755; PMID:10710307
 A:Accession: B81215
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-132 <TET>
 A:Cross-references: GB:AE002386; GB:AE002098; NID:G7225512; PIDN:AAF40742.1; PID:G722551
 A:Experimental source: serogroup B, strain MC58
 R:Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
 Holroyd, S.; Jorgels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream,
 Nature 404, 502-506, 2000
 A>Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
 A:Reference number: A81775; MUID:20222556; PMID:10761919
 A:Accession: G81792
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-132 <PAR>

A:Cross-references: GB:AL162758; GB:AL157959; NID:G7380672; PIDN:CAB85407.1; PID:G738081
 A:Experimental source: serogroup A, strain Z2491
 C:Genetics:
 A:Gene: NMB0291; NMA2196

Query Match 0.6%; Score 7; DB 2; Length 132;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLA 1157

1ue sep 21 14:40:33 2004

Db 45 GGLLLA 51

Search completed: September 21, 2004, 13:42:32
Job time : 32 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:06:29 ; Search time 29 seconds
(without alignments)
2133.081 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188
Sequence: 1 MDLPRGLVVAWALSLWPGFT.....FRSARRRPGLDTPFKYLE 1188

Scoring table: OLIGO
Gapop 60.0 , Capext 60.0

Searched: 141681 seqs, 52070155 residues

Word size : 0

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	911	76.7	1189	1	ITAH_HUMAN
2	15	1.3	285	1	ITAH_CHICK
3	15	1.3	1151	1	ITAH_HUMAN
4	15	1.3	1180	1	ITAH_RAT
5	13	1.1	1167	1	ITAG_HUMAN
6	11	0.9	1170	1	ITA2_BOVIN
7	11	0.9	1181	1	ITA2_HUMAN
8	9	0.8	312	1	CAH4_BOVIN
9	9	0.8	371	1	LEU3_SCHPO
10	9	0.8	1178	1	ITA2_MOUSE
11	8	0.7	78	1	CD24_MOUSE
12	8	0.7	76	1	CD24_RAT
13	8	0.7	80	1	CD24_HUMAN
14	8	0.7	97	1	LST1_HUMAN
15	8	0.7	103	1	FTSB_SALTY
16	8	0.7	105	1	FTSB_KLEAE
17	8	0.7	179	1	IL22_HUMAN
18	8	0.7	234	1	EMBL_CAVPO
19	8	0.7	264	1	COMT_RAT
20	8	0.7	402	1	OPDE_PSEAE
21	8	0.7	423	1	NSMA_HUMAN
22	8	0.7	444	1	ENGA_MYCPE
23	8	0.7	451	1	HISX_COREF
24	8	0.7	467	1	DADR_HUMAN
25	8	0.7	467	1	SIL7_HUMAN
26	8	0.7	470	1	CPBK_MOUSE
27	8	0.7	487	1	XYLC_PSEPU
28	8	0.7	491	1	CPB1_RAT
29	8	0.7	491	1	CPB2_RAT
30	8	0.7	500	1	CPBA_MOUSE
31	8	0.7	604	1	VE1_EPV2
32	8	0.7	605	1	VE1_EPV1
33	8	0.7	811	1	TMS6_HUMAN
					Q9ukx5 homo sapien
					Q90615 gallus gall
					P56199 homo sapien
					P18614 rattus norv
					O75578 homo sapien
					P53710 bos taurus
					P17301 homo sapien
					Q95323 bos taurus
					P18869 schizosacch
					O62469 mus musculu
					P24807 mus musculu
					Q07490 rattus norv
					P25063 homo sapien
					O00453 homo sapien
					O8xep2 salmonella
					Q9eyv3 klebsiella
					O9gxz6 homo sapien
					P35709 cavia porce
					P22734 rattus norv
					Q01602 pseudomonas
					O60906 homo sapien
					O8ewh6 mycoplasma
					O8fnz0 corynebacte
					P21917 homo sapien
					Q9Y286 homo sapien
					Q62397 mus musculu
					P43503 pseudomonas
					P00176 rattus norv
					P04167 rattus norv
					P12791 mus musculu
					P11298 bovine papi
					P03116 bovine papi
					Q6iu80 homo sapien

RESULT 1

ALIGNMENTS

34	0.7	1050	1	ITAS_XENLA
35	0.7	1146	1	ITAI_DROME
36	0.7	1152	1	ITAM_HUMAN
37	0.7	1157	1	DD37_HUMAN
38	0.7	1162	1	ITAD_HUMAN
39	0.7	1163	1	ITAX_HUMAN
40	0.7	1173	1	TSPI_XENLA
41	0.7	1179	1	ITAE_HUMAN
42	0.7	1460	1	PMPC_CHLMU
43	0.7	1507	1	SET2_CABEL
44	0.7	1664	1	RPA1_YEAST
45	0.7	2238	1	RRP1_BUNYM
46	0.6	95	1	LST1_MOUSE
47	0.6	109	1	FEL2_FELCA
48	0.6	110	1	HDEA_ECOLI
49	0.6	110	1	INS1_RAT
50	0.6	110	1	INS_RABIT
51	0.6	139	1	IGF_MXGL
52	0.6	139	1	LAMP_PETMA
53	0.6	141	1	HBAD_CHICK
54	0.6	141	1	HBAD_MELGA
55	0.6	142	1	MX_CHICK
56	0.6	160	1	BIK_HUMAN
57	0.6	160	1	NUOE_AQUAE
58	0.6	161	1	TCH2_ABATH
59	0.6	171	1	YF87_METJA
60	0.6	176	1	YD2F_SCHPO
61	0.6	194	1	HIS7_CHLTE
62	0.6	201	1	Z239_MOUSE
63	0.6	206	1	GPBB_MOUSE
64	0.6	208	1	PSMB_SULSO
65	0.6	213	1	KETA_PYRAE
66	0.6	214	1	CH13_HUMAN
67	0.6	214	1	CYB_CERCE
68	0.6	220	1	FGF3_CHICK
69	0.6	221	1	SDFL_HUMAN
70	0.6	223	1	EMBP_MOUSE
71	0.6	227	1	EMBP_RAT
72	0.6	232	1	EMBL_CAVPO
73	0.6	235	1	YBEU_ECOLI
74	0.6	235	1	YOGA_ECOLI
75	0.6	236	1	RR2_MAIZE
76	0.6	236	1	RR2_ORYSA
77	0.6	236	1	RR2_WHEAT
78	0.6	240	1	IBP6_HUMAN
79	0.6	253	1	GX42_RAT
80	0.6	253	1	Y685_METJA
81	0.6	260	1	COX3_ASTPE
82	0.6	260	1	COX3_PISOC
83	0.6	261	1	COX3_DINSE
84	0.6	261	1	COX3_PARLI
85	0.6	262	1	COX3_BRALA
86	0.6	262	1	COX3_METSE
87	0.6	263	1	LFSA_PASPA
88	0.6	270	1	REP8_HUMAN
89	0.6	276	1	NOSY_PSEST
90	0.6	276	1	UPK_BACSU
91	0.6	295	1	Y234_METJA
92	0.6	301	1	PANE_PYRHO
93	0.6	306	1	COAA_STRMU
94	0.6	309	1	MAT1_HUMAN
95	0.6	309	1	NTCB_SYNPF
96	0.6	311	1	MRAN_CAUCR
97	0.6	311	1	PYDA_ENTFA
98	0.6	312	1	O8H2_HUMAN
99	0.6	313	1	LSU3_NEILA
100	0.6	315	1	CALU_HUMAN

Q06274	xenopus lae
Q24247	drosophila
P11215	homo sapien
O81y37	homo sapien
Q13349	homo sapien
P20702	homo sapien
P35448	xenopus lae
P38570	homo sapien
Q9pjy1	chlamydia m
Q18221	caenorhabdi
P10964	saccharomyc
P20470	bunyamwera
O08843	mus musculu
P30440	felis silve
P26604	escherichia
P01322	rattus norv
P01311	oryctolagus
P22618	myxine glut
P33575	petromyzon
P02001	gallus gall
P81024	meleagris g
P24052	gallus gall
Q13323	homo sapien
O66842	aquifex aeo
P56070	arabidopsis
Q58982	methanococc
Q10263	schizosacch
Q8kef4	chlorobium
P24399	mus musculu
P56400	mus musculu
Q9uxf3	sulfolobus
Oaxsp2	pyrobaculum
Q96ks9	homo sapien
P87419	cerastes ce
P48801	gallus gall
Q9hcn8	homo sapien
Q61878	mus musculu
Q63189	rattus norv
P22032	cavia porce
P77427	escherichia
Q46831	escherichia
P16037	zea mays (m
P12145	oryza sativ
P17933	trititum ae
P24592	homo sapien
Q91xr8	rattus norv
Q58098	methanococc
Q33824	asterina pe
O79552	dinodon sem
P12702	paracentrot
O47425	branchiost
O47491	metridium s
Q05770	pasteurella
O00124	homo sapien
P19845	pseudomonas
Q60258	methanococc
P94507	bacillus su
Q59626	enterococcu
Q8n162	homo sapien
P50180	neisseria l
O43852	homo sapien

ITAH_HUMAN
 ID ITAH_HUMAN STANDARD; PRT; 1189 AA.
 AC Q9UKX5; Q9UKQ1;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-11 precursor.
 GN ITG11.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal heart, and Osteoblast;
 RX MEDLINE=94117678; PubMed=10486209;
 RA Lehnert K., Ni J., Leung E., Gough S.M., Weaver A., Yao W.P., Liu D.,
 RA Wang S.-X., Morris C.M., Krissansen G.W.;
 RT "Cloning, sequence analysis, and chromosomal localization of the novel
 human integrin alpha11 subunit (ITG11).";
 RL Genomics 60:179-187(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Fetal muscle, and Uterus;
 RX MEDLINE=99395147; PubMed=10464311;
 RA Velling T., Kusche-Gullberg M., Seijersen T., Gullberg D.;
 RT "cDNA Cloning and Chromosomal Localization of Human alpha(11)
 Integrin. A collagen-binding, i domain-containing, beta(1)-associated
 integrin alpha-chain present in muscle tissues.";
 RL J. Biol. Chem. 274:25735-25742(1999).
 RN [3]
 RP SEQUENCE OF 954-1188 FROM N.A.
 RC TISSUE=Fibroblast;
 RA Andreu N., Estivill X., Escarceller M., Sumoy L.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INTEGRIN ALPHA-11/BETA-1 IS A RECEPTOR FOR COLLAGEN.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-11
 CC ASSOCIATES WITH BETA-1.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- TISSUE SPECIFICITY: ACCORDING REF.2 HIGHEST LEVELS IN UTERUS AND
 CC HEART. INTERMEDIATE LEVELS IN SKELETAL MUSCLE AND INTERMEDIATE TO
 CC LOW LEVELS IN PANCREAS, KIDNEY AND PLACENTA. ACCORDING REF.1 ALSO
 CC FOUND IN BRAIN, COLON, LUNG, SMALL INTESTINE, STOMACH, TESTIS,
 CC SALIVARY GLANDS, THYROID GLANDS AND PROSTATE. VERY LOW LEVELS IN
 CC PERIPHERAL BLOOD LYMPHOCYTES, FETAL BRAIN AND FETAL LIVER.
 CC -1- DEVELOPMENTAL STAGE: STRONGLY UP-REGULATED IN DIFFERENTIATING
 CC FETAL MUSCLE CELLS (IN VITRO).
 CC -1- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 1 VWFA domain.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF109681; AAF01258.1; -
 CC EMBL; AF137378; AAD51919.2; -
 CC EMBL; AL359064; CAB94392.1; -
 CC HSSP; P17301; IAOX.
 CC
 CC Genew; HGNC:6136; ITG11.
 CC MIM; 604789; -
 CC GO; GO:0008305; C:integrin complex; TAS.
 CC GO; GO:0004895; F:cell adhesion receptor activity; TAS.
 CC GO; GO:0005518; F:collagen binding; TAS.
 CC GO; GO:0007160; P:cell-matrix adhesion; TAS.
 CC GO; GO:0007517; P:muscle development; TAS.
 CC InterPro; IPR000413; Integrin_alpha.

DR InterPro; IPR002035; VWF_A.
 DR Pfam; PF01839; FG-GAP; 3.
 DR Pfam; PF00092; vwa; 1.
 DR PRINTS; PRO1185; INTEGRIN.
 DR PRINTS; PRC0453; VWFADOMAIN.
 DR PRINTS; SM00191; Int_alpha; 5.
 DR SMART; SM00327; vwa; 1.
 DR PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG.
 DR PROSITE; PS0234; VWFA; 1.
 KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 KW Signal; Repeat; Calcium; Magnesium; Polymorphism.
 FT SIGNAL 1 22
 FT CHAIN 23 1189
 FT DOMAIN 23 1142
 FT TRANSMEM 1143 1165
 FT DOMAIN 1166 1189
 FT REPEAT 38 94
 FT REPEAT 102 163
 FT DOMAIN 167 345
 FT REPEAT 359 420
 FT REPEAT 422 475
 FT REPEAT 477 537
 FT REPEAT 539 598
 FT REPEAT 601 653
 FT DOMAIN 1154 1162
 FT DOMAIN 1174 1177
 FT CA_BIND 488 496
 FT CA_BIND 551 559
 FT CA_BIND 613 621
 FT DISULFID 76 83
 FT DISULFID 121 139
 FT DISULFID 129 159
 FT DISULFID 659 668
 FT DISULFID 674 729
 FT DISULFID 781 787
 FT DISULFID 881 893
 FT CARBOHYD 82 82
 FT CARBOHYD 95 95
 FT CARBOHYD 291 291
 FT CARBOHYD 331 331
 FT CARBOHYD 358 358
 FT CARBOHYD 449 449
 FT CARBOHYD 462 462
 FT CARBOHYD 528 528
 FT CARBOHYD 642 642
 FT CARBOHYD 694 694
 FT CARBOHYD 857 857
 FT CARBOHYD 894 894
 FT CARBOHYD 973 973
 FT CARBOHYD 1032 1032
 FT CARBOHYD 1040 1040
 FT VARIANT 433 433
 FT VARIANT 524 524
 FT VARIANT 972 972
 FT VARIANT 1003 1003
 FT VARIANT 1030 1030
 FT VARIANT 1094 1094
 FT VARIANT 133609 MW; 6030308AA4ACD52 CRC64;
 SQ SEQUENCE 1189 AA; 133609 MW; 6030308AA4ACD52 CRC64;
 Query Match 76.7%; Score 911; DB 1; Length 1189;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLPRGLVWAWALSHPGTDTFNMDTRKPRVPGSRTAFFGYTQQCHDSGNKMLVGA 60
 Db 1 MDLPRGLVWAWALSHPGTDTFNMDTRKPRVPGSRTAFFGYTQQCHDSGNKMLVGA 60

Db	901 RAKAVAFELD 911
RESULT 2	
ITAL_CHICK	
ID ITAL_CHICK	STANDARD; PRT; 285 AA.
AC Q30615;	
DT 15-JUL-1998 (Rel. 36, Created)	
DT 15-JUL-1998 (Rel. 36, Last sequence update)	
DT 28-FEB-2003 (Rel. 41, Last annotation update)	
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a)	
DE (Fragment).	
GN ITGA1.	
OC Gallus gallus (Chicken).	
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	

```

QY      164 DIVIVLQSGNSIYEW 178
        |||||
        63 DIVIVLQSGNSIYEW 77
        |||||

RESULT 3
ITAL HUMAN
ID _ITAL HUMAN STANDARD; PRT; 1151 AA.
AC P56199;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-1 (Laminin and collagen receptor) (VLA-1) (CD49a).
GN ITGAL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```


DR	EMBL; X52140; CAA36384.1; -	DR	EMBL; AF074015; AAC31952.1; -
DR	PIR; A35854; A35854	DR	EMBL; AF112345; AAF21944.1; -
DR	PDB; 1CK4; 03-MAY-00	DR	EMBL; AF172723; AAF61638.1; -
DR	InterPro; IPR000413; Integrin_alpha	DR	HSSP; P17301; 1AOK
DR	InterPro; IPR002035; VWF_A	DR	MIM; 604042; -
DR	Pfam; PF01839; FG-GAP; 3	DR	GO; GO:0008305; C:integrin complex; TAS.
DR	Pfam; PF00357; Integrin_A; 1	DR	GO; GO:0004895; F:cell adhesion receptor activity; TAS.
DR	Pfam; PF00092; vwa; 1	DR	GO; GO:0003518; F:collagen binding; TAS.
DR	SMART; SM00191; Int_alpha; 5	DR	GO; GO:0007160; P:cell-matrix adhesion; TAS.
DR	SMART; SM00327; VWA; 1	DR	InterPro; IPR000413; Integrin_alpha
DR	PROSITE; PS00242; INTEGRIN_ALPHA; 1	DR	InterPro; IPR002035; VWF_A
DR	PROSITE; PS02344; VWFA; 1	DR	Pfam; PF01839; FG-GAP; 3
KW	Signal; Repeat; Calcium; Magnesium; 3D-structure	DR	Pfam; PF00092; vwa; 1
FT	SIGNAL 1 28	DR	PRINTS; PR01185; INTEGRINA
FT	CHAIN 29 1180	DR	PRINTS; PR00453; VWFADOMAIN
FT	DOMAIN 23 1142	DR	SMART; SM00191; Int_alpha; 4
FT	TRANSMEM 1143 1165	DR	SMART; SM00327; VWA; 1
FT	DOMAIN 1166 1180	DR	PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG
FT	REPEAT 44 103	DR	PROSITE; PS02344; VWFA; 1
FT	REPEAT ? 7	KW	Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane; Signal; Repeat; Calcium; Magnesium
FT	DOMAIN 175 388		
FT	REPEAT 377 432		
FT	REPEAT 433 484		
FT	REPEAT 485 565		
FT	REPEAT 567 626		
FT	REPEAT 629 681		
FT	CA_BIND 497 505		
FT	CA_BIND 579 587		
FT	CA_BIND 641 649		
FT	SITE 1168 1172		
FT	DISULFID 82 92		
FT	CARBOHYD 100 100		
FT	CARBOHYD 105 105		
FT	CARBOHYD 112 112		
FT	CARBOHYD 217 217		
FT	CARBOHYD 317 317		
FT	CARBOHYD 341 341		
FT	CARBOHYD 402 402		
FT	CARBOHYD 418 418		
FT	CARBOHYD 459 459		
FT	CARBOHYD 531 531		
FT	CARBOHYD 698 698		
FT	CARBOHYD 747 747		
FT	CARBOHYD 779 779		
FT	CARBOHYD 820 820		
FT	CARBOHYD 839 839		
FT	CARBOHYD 882 882		
FT	CARBOHYD 907 907		
FT	CARBOHYD 938 938		
FT	CARBOHYD 965 965		
FT	CARBOHYD 973 973		
FT	CARBOHYD 1007 1007		
FT	CARBOHYD 1084 1084		
FT	CARBOHYD 1103 1103		
FT	CARBOHYD 1114 1114		
SQ	SEQUENCE 1180 AA; 130808 MW; 8E5DA2BE02362EE4 CRC64;		

RESULT 5
ITAG HUMAN STANDARD; PRT; 1167 AA.
AC O75578; Q9UHZ8;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Integrin alpha-10 precursor.
GN ITGA10.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Articular chondrocytes;
RX MEDLINE=98352078; PubMed=9685391;
RA Camper L., Hellman U., Lundgren-Aakerlund E.;
RT "Isolation, cloning, and sequence analysis of the integrin subunit
alpha10, a beta1-associated collagen binding integrin expressed on
chondrocytes";
RL J. Biol. Chem. 273:20389-20389(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Endothelial cells, and Heart;
RX MEDLINE=20169197; PubMed=10702680;
RA Lehnert K., Ni J., Leung E., Gough S.M., Morris C.M., Liu D.,
Wang S.-X., Langley R., Krissansen G.W.;
RT "The integrin alpha10 subunit: expression pattern, partial gene
structure, and chromosomal localization";
RL Cytogenet. Cell Genet. 87:238-244(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-10/BETA-1 IS A RECEPTOR FOR COLLAGEN.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-10
CC -!- ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: Widely expressed with highest expression in
muscle and heart. Found in articular cartilage.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

EMBL; AF074015; AAC31952.1; -
EMBL; AF112345; AAF21944.1; -
EMBL; AF172723; AAF61638.1; -
HSSP; P17301; 1AOK
MIM; 604042; -
GO; GO:0008305; C:integrin complex; TAS.
GO; GO:0004895; F:cell adhesion receptor activity; TAS.
GO; GO:0003518; F:collagen binding; TAS.
GO; GO:0007160; P:cell-matrix adhesion; TAS.
InterPro; IPR000413; Integrin_alpha
InterPro; IPR002035; VWF_A
Pfam; PF01839; FG-GAP; 3
Pfam; PF00092; vwa; 1
PRINTS; PR01185; INTEGRINA
PRINTS; PR00453; VWFADOMAIN
SMART; SM00191; Int_alpha; 4
SMART; SM00327; VWA; 1
PROSITE; PS00242; INTEGRIN_ALPHA; FALSE_NEG
PROSITE; PS02344; VWFA; 1
Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; Repeat; Calcium; Magnesium

Query Match 1.3%; Score 15; DB 1; Length 1180;
Best Local Similarity 100.0%; Fred. No. 5.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
|||||
DB 172 DIVIVLDGNSIYPW 186

FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 1167 INTEGRIN ALPHA-10.
 FT DOMAIN 23 1122 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 1123 1145 POTENTIAL.
 FT DOMAIN 1146 1167 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 38 97 FG-GAP 1.
 FT REPEAT ? ? FG-GAP 2.
 FT DOMAIN 167 350 VWFA.
 FT REPEAT 365 427 FG-GAP 3.
 FT REPEAT 428 482 FG-GAP 4.
 FT REPEAT 483 545 FG-GAP 5.
 FT REPEAT 546 605 FG-GAP 6.
 FT REPEAT 608 660 FG-GAP 7.
 FT DOMAIN 1134 1140 POLY-LEU.
 FT CA_BIND 494 502 POTENTIAL.
 FT CA_BIND 558 566 POTENTIAL.
 FT CA_BIND 620 628 POTENTIAL.
 FT DISULFID 76 86 BY SIMILARITY.
 FT DISULFID 666 675 BY SIMILARITY.
 FT DISULFID 681 736 BY SIMILARITY.
 FT DISULFID 789 795 BY SIMILARITY.
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 336 336 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 364 364 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 839 839 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 921 921 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1011 1011 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 1039 1039 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 844 844 I -> L (IN REF. 2).
 FT CONFLICT 909 909 G -> V (IN REF. 2).
 FT CONFLICT 926 926 E -> D (IN REF. 2).
 SQ SEQUENCE 1167 AA; 127573 MW; AE7D3A1C25C1AEAO CRC64;

Query Match 1.1%; Score 13; DB 1; Length 1167;
 Best Local Similarity 100.0%; Pred. No. 0.00054;
 Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSIYPW 178
 |||||
 DB 169 VIVLDGNSIYPW 181
 |||||

RESULT 6
 ID ITA2_BOVIN STANDARD; PRT; 1170 AA.
 AC P53710;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Integrin alpha-2 precursor (platelet membrane glycoprotein Ia) (GPIa)
 DE (Collagen receptor) (VLA-2 alpha chain) (CD49b) (fragment).
 GN ITGA2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]_TaxID=9913;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94193647; PubMed=7511592;
 RA Kamata T., Puzon W., Takada Y.;
 RT Identification of putative ligand binding sites within I domain of
 RT integrin alpha 2 beta 1 (VLA-2, CD49b/CD29).";
 RL J. Biol. Chem. 269:9659-9663 (1994).
 CC !- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A RECEPTOR FOR LAMININ,
 CC COLLAGEN, COLLAGEN C-PROPEPTIDES, FIBRONECTIN AND E-CADHERIN. IT
 CC RECOGNIZES THE PROLINE-HYDROXYLATED SEQUENCE G-P-P-G-E-R IN
 CC COLLAGEN. IT IS RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER
 CC CELLS TO COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE

CC CC EXPRESSION, FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
 CC CC EXTRACELLULAR MATRIX.
 CC CC !- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
 CC CC ASSOCIATES WITH BETA-1.
 CC CC !- SUBCELLULAR LOCATION: Type I membrane protein.
 CC CC !- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
 CC CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
 CC CC !- SIMILARITY: Belongs to the integrin alpha chain family.
 CC CC !- SIMILARITY: Contains 1 VWFA domain.
 CC CC !- SIMILARITY: Contains 7 FG-GAP repeats.
 CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC CC the European Bioinformatics Institute. There are no restrictions on its
 CC CC use by non-profit institutions as long as its content is in no way
 CC CC modified and this statement is not removed. Usage by and for commercial
 CC CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC CC or send an email to license@isb-sib.ch).
 CC CC -----
 CC CC EMBL; L25886; AAB59255.1; -.
 CC CC PIR; I45914; I45914.
 CC CC HSP; P17301; IAOX.
 CC CC InterPro; IPR000413; Integrin_alpha.
 CC CC InterPro; IPR002035; VWFA.
 CC CC Pfam; PF01839; FG-GAP; 3.
 CC CC Pfam; PF00357; Integrin_A; 1.
 CC CC Pfam; PF00092; vwa; 1.
 CC CC SMART; SM00191; Int_alpha; 5.
 CC CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
 CC CC PROSITE; PS0234; VWFA; 1.
 CC CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 CC CC Platelet; Signal; Repeat; Polymorphism; Calcium; Magnesium.
 CC CC NON_TER 1
 CC CC SIGNAL <1 18 POTENTIAL.
 CC CC CHAIN 19 1170 INTEGRIN ALPHA-2.
 CC CC DOMAIN 19 1121 EXTRACELLULAR (POTENTIAL).
 CC CC TRANSMEM 1122 1143 POTENTIAL.
 CC CC DOMAIN 1144 1170 CYTOPLASMIC (POTENTIAL).
 CC CC REPEAT 34 92 FG-GAP 1.
 CC CC REPEAT ? ? FG-GAP 2.
 CC CC DOMAIN 177 367 VWFA.
 CC CC REPEAT ? ? FG-GAP 3.
 CC CC REPEAT 423 475 FG-GAP 4.
 CC CC REPEAT 477 538 FG-GAP 5.
 CC CC REPEAT 540 599 FG-GAP 6.
 CC CC REPEAT 604 656 FG-GAP 7.
 CC CC CA_BIND 488 496 POTENTIAL.
 CC CC CA_BIND 552 560 POTENTIAL.
 CC CC CA_BIND 616 624 POTENTIAL.
 CC CC SITE 472 474 CELL ATTACHMENT SITE (POTENTIAL).
 CC CC SITE 1146 1150 GFFR MOTIF.
 CC CC DISULFID 72 81 BY SIMILARITY.
 CC CC DISULFID 669 726 BY SIMILARITY.
 CC CC DISULFID 778 784 BY SIMILARITY.
 CC CC DISULFID 854 865 BY SIMILARITY.
 CC CC DISULFID 1008 1039 BY SIMILARITY.
 CC CC DISULFID 1044 1049 BY SIMILARITY.
 CC CC CARBOHYD 94 94 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 101 101 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 332 332 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 421 421 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 449 449 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 464 464 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 688 688 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 748 748 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 945 945 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 1063 1063 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC CARBOHYD 1070 1070 N-LINKED (GLCNAC. .) (POTENTIAL).
 CC CC VARIANT 580 580 G -> V.
 CC CC VARIANT 588 588 R -> K.
 CC CC VARIANT 725 725 R -> S.
 CC CC SEQUENCE 1170 AA; 128929 MW; BECEFLC5F2448FE1 CRC64;

[illegible]

```

FT CARBOHYD 450 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 475 475 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 699 699 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1057 1057 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1074 1074 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1081 1081 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 534 534 K -> E (IN ALLOCANTIGEN HPA-5B;
FT dbSNP:1801106).
FT /FTIG=VAR_003977.
FT TURN 170 171
FT STRAND 173 180
FT TURN 183 184
FT HELIX 188 199
FT TURN 200 201
FT STRAND 204 204
FT TURN 206 207
FT STRAND 209 216
FT TURN 220 224
FT STRAND 226 228
FT HELIX 232 240
FT TURN 241 241
FT HELIX 252 262
FT TURN 263 264
FT HELIX 266 268
FT TURN 269 269
FT STRAND 275 282
FT HELIX 289 291
FT TURN 292 301
FT STRAND 302 303
FT TURN 304 311
FT HELIX 313 317
FT TURN 318 319
FT HELIX 323 330
FT TURN 331 332
FT HELIX 337 340
FT STRAND 341 344
FT HELIX 347 353
FT TURN 354 362
FT TURN 363 363
SQ SEQUENCE 1181 AA; 129235 MW; 7E1B7ED968A94070 CRC64;

```

```

Query Match 0.8%; Score 11; DB 1; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.051;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 495 TDVLVGAPMY 505
Db 506 TDVLVGAPMY 516

```

```

RESULT 8
ID CAH4 BOVIN STANDARD; PRT; 312 AA.
AC Q95323;
DT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Carbonic anhydrase IV precursor (EC 4.2.1.1) (Carbonate dehydratase IV) (CA-IV).
DE IV) (CA-IV).
DE CA4.
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Holstein; TISSUE=Kidney;
RA Tamai S.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBSJ databases.
CC -!- FUNCTION: Reversible hydration of carbon dioxide.
CC -!- CATALYTIC ACTIVITY: H(2)CO(3) = CO(2) + H(2)O.
CC -!- COFACTOR: Zinc.

```

```

CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- SIMILARITY: Belongs to the eukaryotic-type carbonic anhydrase
CC family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U58870; AAB09466.1; -.
CC HSP; P22748; 1ZNC.
CC InterPro; IPR001148; Euk Coanhd.
CC Pfam; PF00194; carb anhydrase; 1.
CC ProDom; PD000865; Euk Coanhd; 1.
CC PROSITE; PS00162; EUK_CO2_ANHYDRASE; FALSE NEG.
KW Glycoprotein; Lipoprotein; GPI-anchor; Membrane; Lyase; Zinc; Signal.
FT SIGNAL 1 18
FT CHAIN 19 284
FT PROPEP 285 312
FT METAL 115 115
FT METAL 117 117
FT METAL 140 140
FT DISULFID 24 36
FT DISULFID 46 229
FT LIPID 284
FT LIPID 284
SQ SEQUENCE 312 AA; 35151 MW; BAE320C09426351 CRC64;

```

```

Query Match 0.8%; Score 9; DB 1; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1154 LLLALLVLA 1162
Db 3 LLLALLVLA 11

```

```

RESULT 9
ID LEU3 SCHPO STANDARD; PRT; 371 AA.
AC P18869;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 3-isopropylmalate dehydrogenase (EC 1.1.1.85) (Beta-IPM dehydrogenase)
DE (IMDH) (3-IPM-DH).
GN LEU1 OR SPBC1A4.02C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Kikuchi Y., Kitazawa Y., Shimatake G., Yamamoto M.;
RA "The primary structure of the leu1+ gene of Schizosaccharomycetes
RA pombe.";
RL Curr. Genet. 14:375-379(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Scouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Felwell T., Fraser A.,
RA Gholroyd S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,

```

RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders R., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Weltjens I., Volckaert E., Aert R., Robben J., Grymonprez B.,
RA Weidmann J., Vansteenkiste E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Porashkin J.,
RA Spakovski G.V., Uesery D., Barrell B.G., Nurse P.,
RT "The genome sequence of *Schizosaccharomyces pombe*."
RL Nature 415:871-880(2002).
CC -!- FUNCTION: Catalyzes the oxidation of 3-carboxy-2-hydroxy-4-
CC methylpentanoate (3-isopropylmalate) to 3-carboxy-4-methyl-2-
CC oxopentanoate. The product decarboxylates to 4-methyl-2-
CC oxopentanoate.
CC -!- CATALYTIC ACTIVITY: 3-carboxy-2-hydroxy-4-methylpentanoate +
CC NAD(+) = 3-carboxy-4-methyl-2-oxopentanoate + NADH.
CC -!- PATHWAY: Leucine biosynthesis; third step.
CC -!- SUBUNIT: Homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the isocitrate and isopropylmalate
CC dehydrogenases family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M36910; AAA35316.1; -.
CC DR EMBL; AL031174; CAA20106.1; -.
CC DR EMBL; T43407; T43407.
CC DR HSSP; P12010; 2AYQ.
CC DR GeneDB_Spombe; SPBC1A4.02c; -.
CC DR InterPro; IPR001804; IsoDH.
CC DR InterPro; IPR004429; LeuB.
CC DR Pfam; PF00180; IsoDH; 1.
CC DR TIGRFAMs; TIGR00169; leuB; 1.
CC DR PROSITE; PS00470; IDH_IMDH; 1.
CC Oxidoreductase; Leucine biosynthesis; NAD.
KW SEQUENCE 371 AA; 39732 MW; 65AA2E6AA94D45EE CRC64;
SQ
Query Match 0.8%; Score 9; DB 1; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 378 DGVLGAVG 386
DB 69 DGVLGAVG 77
RESULT 10
IT2A_MOUSE
ID IT2A_MOUSE STANDARD; PRT; 1178 AA.
AC Q62469; Q62163;
DT 15-JUL-1998 (Rel. 36; Created)
DT 15-JUL-1998 (Rel. 36; Last sequence update)
DT 28-FEB-2003 (Rel. 41; Last annotation update)
DE Integrin alpha-2 precursor (Platelet membrane glycoprotein Ia) (GP1a)
DE (Collagen receptor) (VLA-2 alpha chain) (CD49b).
GN ITGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1] SEQUENCE FROM N.A.
RN STRAIN=C57BL/6 X CBA; TISSUE=Lung;
RC MEDLINE=94363406; PubMed=8081889;
RX Edelman J.M., Chan B.M., Uniyal S., Ondera H., Wang D.Z.,
RA Damjanovich L., Latzer D.B., Finberg R.W., Bergelson J.M.;
RT "The mouse VLA-2 homologue supports collagen and laminin adhesion but
RT not virus binding.";
RT Cell Adhes. Commun. 2:131-143(1994).
[2]
RN SEQUENCE OF 450-1178 FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=94355691; PubMed=7521231;
RA Wu J.E., Santoro S.A.;
RT "Complex patterns of expression suggest extensive roles for the alpha
RT 2 beta 1 integrin in murine development.";
RL Dev. Dyn. 199:292-314(1994).
CC -!- FUNCTION: INTEGRIN ALPHA-2/BETA-1 IS A COLLAGEN RECEPTOR, BEING
CC RESPONSIBLE FOR ADHESION OF PLATELETS AND OTHER CELLS TO
CC COLLAGENS, MODULATION OF COLLAGEN AND COLLAGENASE GENE EXPRESSION,
CC FORCE GENERATION AND ORGANIZATION OF NEWLY SYNTHESIZED
CC EXTRACELLULAR MATRIX. IT IS ALSO A RECEPTOR FOR LAMININS, COLLAGEN
CC C-PROPEPTIDES AND E-CADHERIN. MICE HOMOZYGOUS FOR A NULL MUTATION
CC IN THE ALPHA-2 DIE VERY EARLY IN EMBRYOGENESIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-2
CC ASSOCIATES WITH BETA-1.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWF A DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWF A domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z29987; CAA82877.1; -.
CC DR EMBL; X75427; CAA53178.1; -.
CC DR FIR; S44142; S44142.
CC DR HSSP; P17301; IAOX.
CC DR MGD; MGI:96600; Itga2.
CC DR InterPro; IPR000413; Integrin_alpha.
CC DR InterPro; IPR002035; VWF A.
CC DR Pfam; PF01839; FG-GAP; 3.
CC DR Pfam; PF00357; integrin_A; 1.
CC DR Pfam; PF00092; vwa; 1.
CC DR Pfam; SM00191; Int_alpha; 4.
CC DR SMART; SM00327; VWA; 1.
CC DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC DR PROSITE; PS0234; VWF A; 1.
CC Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Platelet; Signal; Repeat; Calcium; Magnesium.
FT SIGNAL 1 26
FT CHAIN 27 1178 INTEGRIN ALPHA-2.
FT DOMAIN 27 1129 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 1130 1151 POTENTIAL.
FT DOMAIN 1152 1178 CYTOPLASMIC (POTENTIAL).
FT REPEAT 42 100 FG-GAP 1.
FT REPEAT ? 2 FG-GAP 2.
FT DOMAIN 185 375 VWF A.
FT REPEAT ? ? FG-GAP 3.
FT REPEAT 431 463 FG-GAP 4.
FT REPEAT 485 546 FG-GAP 5.
FT REPEAT 548 607 FG-GAP 6.
FT REPEAT 612 664 FG-GAP 7.
FT CA_BIND 496 POTENTIAL.

```

FT CA_BIND 560 568 POTENTIAL.
FT CA_BIND 624 632 POTENTIAL.
FT SITE 480 482 CELL ATTACHMENT SITE (POTENTIAL).
FT SITE 1154 1158 OFFKX MOTIF.
FT DISULFID 80 89 BY SIMILARITY.
FT DISULFID 677 734 BY SIMILARITY.
FT DISULFID 786 792 BY SIMILARITY.
FT DISULFID 862 873 BY SIMILARITY.
FT DISULFID 1016 1047 BY SIMILARITY.
FT DISULFID 1052 1057 BY SIMILARITY.
FT CARBOHYD 102 102 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 109 109 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 429 429 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 457 457 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 472 472 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1054 1054 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1071 1071 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1078 1078 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1178 AA; 128926 MW; 1F194B9C0240F465 CRC64;

Query Match 0.8%; Score 9; DB 1; Length 1178;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAP 503
DB 503 TDVLLVGAP 511

RESULT 11
CD24 MOUSE
ID CD24 MOUSE STANDARD; PRT; 76 AA.
AC P24807; P26691;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor (M1/69-J11D heat stable antigen)
DE (HSA) (Nectadrin) (IY-52) (X62 heat stable antigen) (R13-AG).
DE CD24 OR CD24A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90361906; PubMed=2118158;
RA Kay R., Takei F., Humphries R.K.;
RT "Expression cloning of a cDNA encoding M1/69-J11D heat-stable
antigens";
RL J. Immunol. 145:1952-1959 (1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6 X CBA, and Swiss albino X BALB/c; TISSUE=Spleen;
RC MEDLINE=91209380; PubMed=2019286;
RA Wanger R.H., Avane M., Bose R., Koehler G., Nielsen P.J.;
RT "The genes for a mouse hematopoietic differentiation marker called
the heat-stable antigen.";
RL Eur. J. Immunol. 21:1039-1046 (1991).
RN [3]
RP REVISIONS.
RA Nielsen P.J.;
RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6 X CBA; TISSUE=Spleen;
RC MEDLINE=94043127; PubMed=8226859;
RA Wanger R.H., Rochelle J.M., Seidin M.F., Koehler G., Nielsen P.J.;
RT "The heat stable antigen (mouse CD24) gene is differentially
regulated but has a housekeeping promoter.";
RL J. Biol. Chem. 268:23345-23352 (1993).
RN [5]
RP SEQUENCE OF 27-53.

```

```

RC STRAIN=C57BL/6;
RX MEDLINE=92412120; PubMed=1530634;
RA Hitsumoto Y., Nakano A., Ohnishi H., Hamada F., Saheki S.,
RA Takeuchi N.;
RT "Purification of the murine heat-stable antigen from erythrocytes.";
RL Biochem. Biophys. Res. Commun. 187:773-777 (1992).
CC -!- FUNCTION: May have a specific role to play in early thymocyte
development.
CC -!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -!- TISSUE SPECIFICITY: In lymphoid, myeloid, and erythroid cells.
CC -!- PTM: Extensively O-glycosylated (By similarity).
CC -!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
DR EMBL; M58661; AAA39481.1; -
DR EMBL; X56469; CAA39841.1; -
DR EMBL; X72910; CAA51415.1; -
DR EMBL; X53825; CAA37822.1; -
DR PIR; A43537; A43537
DR MGI; MGI:88323; Cd24a.
KW Glycoprotein; GPI-anchor; Membrane; Signal; Antigen; Lipoprotein.
FT SIGNAL 1 26
FT CHAIN 27 53
FT PROPEP 54 76 SIGNAL TRANSDUCER CD24.
FT CARBOHYD 27 27 REMOVED IN MATURE FORM (POTENTIAL).
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .).
FT CARBOHYD 39 39 O-LINKED (PROBABLE).
FT CARBOHYD 48 48 N-LINKED (GLCNAC. . .).
FT LIPID 53 53 N-LINKED (GLCNAC. . .).
SQ SEQUENCE 76 AA; 7797 MW; 6853F121B33625EB CRC64;

Query Match 0.7%; Score 8; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 12 GLLLLALL 19

RESULT 12
CD24 RAT
ID CD24 RAT STANDARD; PRT; 76 AA.
AC Q07490;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor (Heat stable antigen) (HSA)
DE (Nectadrin).
GN CD24A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Embryonic brain;
RX MEDLINE=94122434; PubMed=8292828;
RA Shirasawa T., Akashi T., Sakamoto K., Takahashi H., Maruyama N.,
RA Hirokawa K.;
RT "Gene expression of CD24 core peptide molecule in developing brain
and developing non-neural tissues.";
RL Dev. Dyn. 198:1-13 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Fischer;

```

MEDLINE=97157759; PubMed=9004036;
Magnaldo T.A., Barrandon Y.;
CD24 (heat stable antigen, nectadin), a novel keratinocyte
differentiation marker, is preferentially expressed in areas of the
hair follicle containing the colony-forming cells.";
J. Cell Sci. 109:3035-3045(1996).
CC CC CC

-!- FUNCTION: May have a pivotal role in cell differentiation. The
triggering mechanism of signal transduction may be due to the
interactions of differentiating cells with the matrix substrate
via the carbohydrate structure of the molecule. In this way, the
signal transducer can play very different roles in different cell
types as a direct consequence of its glycosylation.
CC CC CC

-!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC CC CC

-!- TISSUE SPECIFICITY: Expressed in the central nervous system, in
postmitotic cells of spinal cord, hindbrain, midbrain and
forebrain. Expressed in epithelium during the development of non-
neural tissues. Expressed in tooth development, specifically in
mesenchymal cells differentiating into odontoblast in dental
papilla, as well as in the developing eye and hair follicle.
CC CC CC

-!- DEVELOPMENTAL STAGE: Detected in primitive ectoderm, mesoderm and
ventral endoderm; down-regulated when organogenesis is completed.
CC CC CC

-!- PTM: Extensively O-glycosylated (By similarity). The carbohydrate
structure may be regulated in a tissue-specific and developmental
stage-specific manner.
CC CC CC

-!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; Z11863; CAA77731.1; --
EMBL; U49062; AAA91470.1; --
PIR; I53107; I53107.
Glycoprotein; GPI-anchor; Membrane; Signal; Differentiation;
Lipoprotein.
SIGNAL 1 26 POTENTIAL.
CHAIN 27 56 SIGNAL TRANSDUCER CD24.
PROPEP 57 76 REMOVED IN MATURE FORM (BY SIMILARITY).
CARBOHYD 27 27 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 37 37 N-LINKED (GLCNAC. .) (POTENTIAL).
CARBOHYD 48 48 N-LINKED (GLCNAC. .) (POTENTIAL).
LIPID 56 56 GPI-anchor amidated serine (Potential).
SEQUENCE 76 AA; 7862 MW; 42846E70EC39D958 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 76;
Best Local Similarity 100.0%; Pred.No. 4.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLALL 1159
| | | | | |
Db 12 GLLLALL 19

RESULT 13
CD24 HUMAN
ID CD24_HUMAN STANDARD; PRT; 80 AA.
AC P25063; Q16257;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Signal transducer CD24 precursor.
DE CD24 OR CD24A.
GN GN Homo sapiens (Human).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
ON NCBI_TaxID=9606;
RX [1]
RP MEDLINE=91332458; PubMed=1831224;
SEQUENCE FROM N.A.

Kay R., Rosten P.M., Humphries R.K.;
CD24, a signal transducer modulating B cell activation responses, is
very short peptide with a glycosyl phosphatidylinositol membrane
anchor."
J. Immunol. 147:1412-1416(1991).
[2]
SEQUENCE FROM N.A., AND VARIANT VAL-57.
MEDLINE=93007871; PubMed=1327504;
Jackson D., Waibel R., Weber E., Ball J., Stahel R.A.;
CD24, a signal-transducing molecule expressed on human B cells, is a
major surface antigen on small cell lung carcinomas."
Cancer Res. 52:5264-5270(1992).
[3]
SEQUENCE FROM N.A.
TISSUE-Ovary:
MEDLINE=42386257; Pubmed=12477932;
Straussberg R.L., Feinsold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
Alschul S.F., Zeenberg B., Buettow K.H., Schaefer C.P., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huilyk S.W.,
Vallalun D.K., Muszy D.M., Sedergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Sherchenko V., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
Schmerch A., Schein J.E., Jones J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[4]
SEQUENCE OF 1-76 FROM N.A.
MEDLINE=95048364; PubMed=7959762;
Hough M.R., Rosten P.M., Sexton T.L., Kay R., Humphries R.K.;
Mapping of CD24 and homologous sequences to multiple chromosomal
loci."
Genomics 22:154-161(1994).
-!- FUNCTION: Modulates B-cell activation responses. Signaling could
be triggered by the binding of a lectin-like ligand to the CD24
carbohydrates, and transduced by the release of second messengers
derived from the GPI-anchor. Promotes AG-dependent proliferation
of B-cells, and prevents their terminal differentiation into
antibody-forming cells.
-!- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
-!- TISSUE SPECIFICITY: B-cells.
-!- INDUCTION: Expression is lost when primary B-cells are induced to
differentiate in antibody-forming cells.
-!- PTM: Extensively O-glycosylated.
-!- SIMILARITY: TO OTHER MAMMALIAN SPECIES CD24.
-!- DATABASE: NAME=PROW; NOTE=cd guide CD24 entry;
WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd24.htm".
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; M58664; AAA35665.1; --
EMBL; X59397; CAA49195.1; --
EMBL; I33930; AAB58807.1; --
EMBL; BC007674; AAH07674.1; --
EMBL; S75311; AAD14170.1; ALT_INIT.
PIR; I56114; A48996.
GENEW; HGNC:1645; CD24.

DR MIM: 600074; -- P:humoral immune response; TAS.
 DR GO:0006959; Gpi-anchor; Membrane; Signal; Polymorphism; Lipoprotein.
 KW Glycoprotein; Gpi-anchor; Membrane; Signal; Polymorphism; Lipoprotein.
 FT SIGNAL 1 26
 FT CHAIN 27 59
 FT PROPEP 60 80
 FT CARBOHYD 36 36
 FT CARBOHYD 52 52
 FT LIPID 59 59
 FT VARIANT 57 57
 FT CONFLICT 12 12
 FT CONFLICT 44 44
 FT CONFLICT 44 44
 SQ SEQUENCE 80 AA; 8083 MW; DB1988B680F833F CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 80;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1152 GLLLLALL 1159
 DB 12 GLLLLALL 19
 RESULT 14
 LST1_HUMAN STANDARD; PRT; 97 AA.
 ID LST1_HUMAN
 AC O00453; O00452; O00454; Q13669; Q9UUR5; Q9UUR6; Q9UUR7; Q9UUR8;
 AC Q9UUR5;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Leukocyte specific transcript 1 protein (B144 protein).
 GN LST1 OR B144
 OS Homo sapiens (Human)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A. (ISOFORM 3), TISSUE SPECIFICITY, AND INDUCTION.
 RC TISSUE=Blood;
 RX MEDLINE=86006565; PubMed=7590964;
 RA Holzinger I., de Baey A., Messer G., Kiek G., Zwierzina H.,
 RA Weiss E.H.;
 RA "Cloning and genomic characterization of LST1: a new gene in the human
 TNF region.";
 RL Immunogenetics 42:315-322(1995).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 4), TISSUE SPECIFICITY, AND
 INDUCTION.
 RC TISSUE=Lymphoblast;
 RX MEDLINE=98035883; PubMed=9367684;
 RA de Baey A., Fellerhoff B., Maier S., Martinozzi S., Weidle U.,
 RA Weiss E.H.;
 RA "Complex expression pattern of the TNF region gene LST1 through
 differential regulation, initiation, and alternative splicing.";
 RL Genomics 45:591-600(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORMS 5; 6; 7; 8 AND 9), FUNCTION, AND
 SUBCELLULAR LOCATION.
 RC TISSUE=Peripheral blood;
 RX MEDLINE=20171517; PubMed=10706707;
 RA Rollinger-Holzinger I., Sibl B., Pauly M., Griesser U., Hentges F.,
 RA Auer B., Pall G., Schratzberger P., Niederwieser D., Weiss E.H.,
 RA Zwierzina H.;
 RA "LST1: a gene with extensive alternative splicing and immunomodulatory
 function.";
 RL J. Immunol. 164:3169-3176(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
 RA Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
 RA Lasky S., Hood L.;

RT "Sequence of the human major histocompatibility complex class III
 region.";
 RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Hirakawa M., Yamaguchi H., Imai K., Shimada J.;
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP FUNCTION, AND SUBCELLULAR LOCATION.
 RX MEDLINE=21372017; PubMed=11478849;
 RA Ragunathan A., Sivakamasundari R., Wolenski J., Poddar R.,
 RA Weissman M.;
 RL "Functional analysis of B144/LST1: a gene in the tumor necrosis
 factor cluster that induces formation of long filopodia in eukaryotic
 cells.";
 RT Exp. Cell Res. 268:230-244(2001).
 RL CC
 CC -!- FUNCTION: Possible role in modulating immune responses. Isoforms 1
 and 2 have an inhibitory effect on lymphocyte proliferation.
 CC Induces morphological changes including production of filopodia
 and microspikes when overexpressed in a variety of cell types and
 may be involved in dendritic cell maturation.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Also detected in
 a perinuclear region corresponding to the localization of the
 Golgi apparatus and throughout the cytoplasm.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=9;
 CC Comment=Additional isoforms seem to exist;
 CC Name=1; Synonyms=LST1/A;
 CC IsoId=O00453-1; Sequence=Displayed;
 CC Name=2; Synonyms=LST1/C;
 CC IsoId=O00453-2; Sequence=VSP_050578;
 CC Name=3; Synonyms=plst1;
 CC IsoId=O00453-3; Sequence=VSP_050579;
 CC Name=4; Synonyms=LST1/E;
 CC IsoId=O00453-4; Sequence=VSP_050577;
 CC Name=5; Synonyms=LST1/K;
 CC IsoId=O00453-5; Sequence=VSP_050584;
 CC Name=6; Synonyms=LST1/L;
 CC IsoId=O00453-6; Sequence=VSP_050583, VSP_050587;
 CC Name=7; Synonyms=LST1/J;
 CC IsoId=O00453-7; Sequence=VSP_050583, VSP_050586;
 CC Name=8; Synonyms=LST1/M;
 CC IsoId=O00453-8; Sequence=VSP_050580;
 CC Name=9; Synonyms=LST1/N;
 CC IsoId=O00453-9; Sequence=VSP_050581, VSP_050582;
 CC -!- TISSUE SPECIFICITY: Expressed in lung, tonsil, thymus, placenta,
 kidney, fetal spleen, fetal liver and brain.
 CC -!- INDUCTION: By interferon gamma.
 CC -!- SIMILARITY: Belongs to the LST1 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 DR EMBL; U00921; AAB57724.1; --
 DR EMBL; AF000424; AAB86998.1; --
 DR EMBL; AF000425; AAB86999.1; --
 DR EMBL; AF000426; AAB87000.1; --
 DR EMBL; Y18486; CAB59904.1; --
 DR EMBL; Y18487; CAB59905.1; --
 DR EMBL; Y18488; CAB60038.1; --
 DR EMBL; Y18489; CAB59906.1; --
 DR EMBL; Y18490; CAB59903.1; --
 DR EMBL; AF129756; AAD18090.1; --
 DR EMBL; AF000505; BAB63394.1; --
 DR Genew; HGNC:14189; LST1.
 DR MIN; 109170; --
 DR GO; GO:0006955; P:immune response; NAS.
 DR GO; GO:0009653; P:morphogenesis; NAS.

InterPro; IPR007775; LST1.
Pfam; PF05083; LST1; 1.
Immune response; Cell shape;
TRANSMEM 8 28
VARSPPLIC 1 46

MLSRNDICFYGGIGLGGILLAVLLSACLWLHRRVKRL
ERSWT -> MIVYSTGAWGWAGSCFWQSFPCPACVGCIEE
HLSSWSQA (in isoform 2).
/FTid=VSP_050578
MLSRNDICFYGGIGLGGILLAVLLSACLWLHRRVKRL
ERSWT -> MIVYSTGAWGWAGSCFWQSFPCPACVGCIEE
EA (in isoform 3).
/FTid=VSP_050579
VARSPPLIC 1 45
MLSNDICFYGGIGLGGILLAVLLSACLWLHRRVKRL
ERSW -> M (in isoform 4).
/FTid=VSP_050577
DICYGGIGLGGILLAVLLSACLWLHRRVKRLERSWTQ
GSSEQLHYASLQRLPVPSSSGPDLRGRKEDPRADY
ACTAENKPT -> APSVLVEPGGLLRAGTPLCISAEASAQ
Q (in isoform 8).
/FTid=VSP_050580
DICIVGG -> EASAAQQ (in isoform 9).
/FTid=VSP_050581
Missing (in isoform 9).
/FTid=VSP_050582
Missing (in isoform 6 and isoform 7).
/FTid=VSP_050583
VKRLERSWTQGSSEQLHYASLQRLPVPSSSGPDLRGRDKR
CTKEDPRADYACTAENKPT -> GPGLLRAGTPLCISAEEA
SAQQ (in isoform 5).
/FTid=VSP_050584
TQGSSEQLHYASLQRLPVPSSSGPDLRGRDKGTGEDPRA
DYACTAENKPT -> HSLSVLPFGGLLRAGTPLCISAEEA
AQO (in isoform 7).
/FTid=VSP_050586
Missing (in isoform 6).
/FTid=VSP_050587
AAO3C761E787AF94 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 97;
Best Local Similarity 100.0%; Pred.No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1150 LGLLLLLA 1157
|||||||
Db 16 LGLLLLLA 23

RESULT 15
FTSB_SALTY STANDARD; PRT; 103 AA.

ID FTSB_SALTY STANDARD; PRT; 103 AA.
AC Q8XEF2;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
DE Cell division protein ftsB homolog.
GN FTSB OR STM2931 OR STY3056 OR T2832.
OS Salmonella typhimurium, and
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601,
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=S.typhimurium; STRAIN=IT2 / SGSCI412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.B., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium LT2.";

Nature 413:852-856(2001).

[2]
SEQUENCE FROM N.A.
SPECIES=S.typhi; STRAIN=CT18;
MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
Churcher C., Mungall K.L., Bentley S.D., Holden M.T.G., Sebahia M.,
Baker S., Basham D., Brooks K., Chillingworth T., Connor P.,
Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
Krogh A., Larsen R.S., Leather S., Moule S., O'Gaora P., Parry C.,
Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella enterica serovar Typhi CT18";
Nature 413:848-852(2001).

[3]
SEQUENCE FROM N.A.
SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
MEDLINE=22531367; PubMed=12644504;
RX Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
Burkand V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2 and Ct18";
J. Bacteriol. 185:2330-2337(2003).

CC -!- FUNCTION: Required for the cell division process (By similarity).
CC -!- SUBUNIT: May interact withftsL (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane (Potential). Colocalizes withftsLto the division site (By similarity).

CC -!- SIMILARITY: Belongs to theftsBfamily.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch).

DR EMBL; AE008833; AAL21811.1; -
DR EMBL; AL627276; CAD06037.1; -
DR EMBL; AE016843; AAO70389.1; -
DR StvGene; SG?22??; ftsB.
DR HAMAP; MF_00599; -. 1.
DR InterPro; IPR007060; DivIC.
DR Pfam; PF04977; DivIC; 1.
KW Cell division; Transmembrane; Inner membrane; Coiled coil;
KW Complete proteome.

FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 4 21 POTENTIAL.
FT DOMAIN 22 103 PERIPLASMIC (POTENTIAL).
FT DOMAIN 30 74 COILED COIL (POTENTIAL).
SQ SEQUENCE 103 AA; 11575 MW; F3B559AB77E8ACEB CRC64;

Query Match 0.7%; Score 8; DB 1; Length 103;
Best Local Similarity 100.0%; Pred.No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1153 LLLLALLV 1160
|||||||
Db 6 LLLLALLV 13

RESULT 16
FTSB_KLEAF STANDARD; PRT; 105 AA.
ID FTSB_KLEAF STANDARD; PRT; 105 AA.
AC Q9EYU3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division protein ftsB homolog.
GN FTSB.

InterPro; IPR007775; LST1.
Pfam; PF05083; LST1; 1.
Immune response; Cell shape;
TRANSMEM 8 28
VARSPPLIC 1 46
MLSRNDICFYGGGGLGGLLLAVLLSACLWLHRRVKRL
HLLSMSQA -> MIVYSTGAWGAGSCFWQSFPCPACVGCIEE
/FTid=VSP 050578
MLSRNDICFYGGGGLGGLLLAVLLSACLWLHRRVKRL
ERSWT -> MIVYSTGAWGAGSCFWQSFPCPACVGCIEE
EA (in isoform 3).
/FTid=VSP 050579
VARSPPLIC 1 45
MLSRNDICFYGGGGLGGLLLAVLLSACLWLHRRVKRL
ERSW -> M (in isoform 4).
/FTid=VSP 050577
DICIVGGGGLGGLLLAVLLSACLWLHRRVKRLERSWTO
GSSQELHYASLQRLVPSSGPDRLGRDKRGKEDPRADY
ACTAENKPT -> APSVLVEGGLLRAGTPLCISAEASQAQ
Q (in isoform 8).
/FTid=VSP 050580
DICIVGG -> EASAAQ (in isoform 9).
/FTid=VSP 050581
Missing (in isoform 9).
/FTid=VSP 050582
Missing (in isoform 6 and isoform 7).
/FTid=VSP 050583
VKRLERSWTGSSQELHYASLQRLVPSSGPDRLGRDKR
CTKEDPRADYACTAENKPT -> GGLLRAGTPLCISAEAA
SAQQ (in isoform 5).
/FTid=VSP 050584
TQSSQELHYASLQRLVPSSGPDRLGRDKRGKEDPRADY
DYACTAENKPT -> HLSVLVPGGLLRAGTPLCISAEAA
AQ (in isoform 7).
/FTid=VSP 050586
Missing (in isoform 6).
/FTid=VSP 050587
AA03C761E787AF94 CRC64;
SEQUENCE 97 AA; 10822 MW; 10822 MW;
Query Match 0.7%; Score 8; DB 1; Length 97;
Best Local Similarity 100.0%; Pred. No. 6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1150 IGGLLLLL 1157
| | | | | | | | | |
Db 16 LGGLLLLA 23
RESULT 15
FTSB_SALTY STANDARD; PRT; 103 AA.
ID FTSB_SALTY STANDARD; PRT; 103 AA.
AC Q8XEF2;
DR 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 10-OCT-2003 (Rel. 42, Last annotation update)
Cell division protein ftsB homolog.
FTSB OR STM2931 OR STY3056 OR T2832.
OS Salmonella typhimurium, and
Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium; STRAIN=IT2 / SGSC1412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.B., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
Nature 413:852-856(2001).
[2]
SEQUENCE FROM N.A.
SPECIES=S.typhi; STRAIN=CT18;
MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Houlden M.T.G., Sebahia M.,
Churcher C., Mungall K.L., Bentley S.D., Chillingworth T., Connor P.,
Baker S., Basham D., Brooks K., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
Krogh A., Larsen R.S., Leather S., Moule S., O'Garra P., Parry C.,
Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
Nature 413:848-852(2001).
[3]
SEQUENCE FROM N.A.
SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
Burkand V., Kodyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Required for the cell division process (By similarity).
CC -!- SUBUNIT: May interact with ftsL (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane
(Potential). Colocalizes with ftsL to the division site (By
similarity).
CC -!- SIMILARITY: Belongs to the ftsB family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation on
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch.
CC EMBL; AE008833; AAL21811.1; -
DR EMBL; AL627276; CAD06037.1; -
DR EMBL; AE016843; AAO70389.1; -
DR StvGene; SG22222; ftsB.
DR HAMAP; MF_00599; -; 1.
DR InterPro; IPR007060; DivIC.
DR Pfam; PF04977; DivIC; 1.
KW Cell division; Transmembrane; Inner membrane; Coiled coil;
KW Complete proteome.
FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 4 21 POTENTIAL.
FT DOMAIN 22 103 PERIPLASMIC (POTENTIAL).
FT DOMAIN 30 74 COILED COIL (POTENTIAL).
SQ SEQUENCE 103 AA; 11575 MW; F35559AB77E8ACEB CRC64;
Query Match 0.7%; Score 8; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1153 LLLLLLV 1160
| | | | | | | | | |
Db 6 LLLLLLV 13
RESULT 16
FTSB_KLEAF STANDARD; PRT; 105 AA.
ID FTSB_KLEAF STANDARD; PRT; 105 AA.
AC Q9EYV3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division protein ftsB homolog.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602, 601;
RN [1]
SEQUENCE FROM N.A.
SPECIES=S.typhimurium; STRAIN=IT2 / SGSC1412 / ATCC 700720;
RC MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.B., Spieth J., Clifton S.W., Latreille P.,
Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
Waterston R., Wilson R.K.;
RT "Complete genome sequence of Salmonella enterica serovar Typhimurium
LT2.";
Nature 413:852-856(2001).
[2]
SEQUENCE FROM N.A.
SPECIES=S.typhi; STRAIN=CT18;
MEDLINE=21534947; PubMed=11677608;
RA Parkhill J., Dougan G., James K.D., Houlden M.T.G., Sebahia M.,
Churcher C., Mungall K.L., Bentley S.D., Chillingworth T., Connor P.,
Baker S., Basham D., Brooks K., Davies R.M., Dowd L., White N., Farrar J.,
Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
Krogh A., Larsen R.S., Leather S., Moule S., O'Garra P., Parry C.,
Quail M.A., Rutherford K., Simmonds M., Skelton J., Stevens K.,
Whitehead S., Barrall B.G.;
RT "Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18.";
Nature 413:848-852(2001).
[3]
SEQUENCE FROM N.A.
SPECIES=S.typhi; STRAIN=Ty2 / ATCC 700931;
MEDLINE=22531367; PubMed=12644504;
RA Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
Burkand V., Kodyanni V., Schwartz D.C., Blattner F.R.;
RT "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
and CT18.";
J. Bacteriol. 185:2330-2337(2003).
CC -!- FUNCTION: Required for the cell division process (By similarity).
CC -!- SUBUNIT: May interact with ftsL (By similarity).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane
(Potential). Colocalizes with ftsL to the division site (By
similarity).
CC -!- SIMILARITY: Belongs to the ftsB family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation on
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
or send an email to license@isb-sib.ch.
CC EMBL; AE008833; AAL21811.1; -
DR EMBL; AL627276; CAD06037.1; -
DR EMBL; AE016843; AAO70389.1; -
DR StvGene; SG22222; ftsB.
DR HAMAP; MF_00599; -; 1.
DR InterPro; IPR007060; DivIC.
DR Pfam; PF04977; DivIC; 1.
KW Cell division; Transmembrane; Inner membrane; Coiled coil;
KW Complete proteome.
FT DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 4 21 POTENTIAL.
FT DOMAIN 22 103 PERIPLASMIC (POTENTIAL).
FT DOMAIN 30 74 COILED COIL (POTENTIAL).
SQ SEQUENCE 103 AA; 11575 MW; F35559AB77E8ACEB CRC64;
Query Match 0.7%; Score 8; DB 1; Length 103;
Best Local Similarity 100.0%; Pred. No. 6.3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1153 LLLLLLV 1160
| | | | | | | | | |
Db 6 LLLLLLV 13
RESULT 16
FTSB_KLEAF STANDARD; PRT; 105 AA.
ID FTSB_KLEAF STANDARD; PRT; 105 AA.
AC Q9EYV3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cell division protein ftsB homolog.
OS Salmonella typhi.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Salmonella.
OX NCBI_TaxID=602,

OS Klebsiella aerogenes.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Klebsiella.
 OX NCBI_TaxID=28451;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=W70;
 RX MEDLINE=20566700; PubMed=11114933;
 RA Koiko M.M., Kapetanovich L.A., Lawrence J.G.;
 RT "Alternative pathways for sirtone synthesis in Klebsiella
 aerogenes.";
 RL J. Bacteriol. 183:328-335(2001).
 CC -!- FUNCTION: Required for the cell division process (By similarity).
 CC -!- SUBUNIT: May interact with ftsL (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Inner membrane
 CC (Potential). Colocalizes with ftsL to the division site (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the ftsB family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AF308468; AAG42461.1; -
 CC HANAP; MF 00599; -; 1 -
 CC InterPro; IPR007060; DivIC.
 CC Pfam; PF04977; DivIC; 1.
 CC Cell division; Transmembrane; Inner membrane; Coiled coil.
 KW DOMAIN 1 3 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 4 21 POTENTIAL.
 FT DOMAIN 22 105 PERIPLASMIC (POTENTIAL).
 FT DOMAIN 31 74 COILED COIL (POTENTIAL).
 SQ SEQUENCE 105 AA; 11951 MW; 530471363FD3112A CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 105;
 Best Local Similarity 100.0%; Pred. No. 6.4;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1153 LLLALLLV 1160
 Db 6 LLLALLLV 13
 RESULT 17
 IL22_HUMAN
 ID IL22_HUMAN STANDARD; PRT; 179 AA.
 AC Q9GZK6; 2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Interleukin-22 precursor (IL-22) (IL-10-related T-cell-derived
 DE inducible factor) (IL-TIF) (UNQ3099/PRO10096).
 GN IL22 OR ILTIF.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20420346; PubMed=10954742;
 RA Dumontier L., Van Roost E., Coliau D., Renaud J.-C.;
 RT "Human interleukin-10-related T cell-derived inducible factor:
 RT molecular cloning and functional characterization as an hepatocyte-
 RT stimulating factor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:10144-10149(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21069354; PubMed=11197690;
 RA Dumontier L., Van Roost E., Coliau D., Ameys G., Michaux L.,

RA Renaud J.-C.;
 RT "IL-TIF/IL-22: genomic organization and mapping of the human and mouse
 RT genes.";
 RL Genes Immun. 1:488-494(2000).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20469498; PubMed=10875937;
 RA Xie M.-H., Aggarwal S., Ho W.-H., Foster J., Zhang Z., Stinson J.,
 RA Wood W.I., Goddard A.D., Gurney A.L.;
 RT "Interleukin (IL)-22, a novel human cytokine that signals through the
 RT interferon receptor-related proteins CRP2-4 and IL-22R.";
 RL J. Biol. Chem. 275:31335-31339(2000).
 RN [4]
 RP SEQUENCE FROM N.A., AND VARIANT GLY-158.
 RA Rieder M.J., Carrington D.P., Chung M.-W., Lee K.L., Poel C.L., Yi Q.,
 RA Nickerson D.A.;
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22887296; PubMed=12975309;
 RA Clark H.F., Gurney A.L., Abaya E., Baker K., Baldwin D., Brush J.,
 RA Chen J., Chow B., Chui C., Crowley C., Currell B., Deuel B., Dowd P.,
 RA Eaton D., Foster J., Grimaldi C., Gu Q., Hass P.E., Heldens S.,
 RA Huang A., Kim H.S., Klimowski L., Jin Y., Johnson S., Lee J.,
 RA Lewis L., Liao D., Mark M., Robbie E., Sanchez C., Schoenfeld J.,
 RA Seehagiri S., Simmons L., Smith J., Smith V., Stinson J., Vagts A.,
 RA Vandlen R., Watanabe C., Wiedand D., Woods K., Xie M.-H., Yansura D.,
 RA Yi S., Yu G., Yuan J., Zhang M., Zhang Z., Goddard A., Wood W.I.,
 RA Godowski P.;
 RT "The secreted protein discovery initiative (SPDI), a large-scale
 RT effort to identify novel human secreted and transmembrane proteins: a
 RT bioinformatics assessment.";
 RL Genome Res. 13:2265-2270(2003).
 CC -!- FUNCTION: Cytokine that contributes to the inflammatory response
 CC in vivo.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: Belongs to the IL-10 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ277247; CAC06085.1; -
 CC EMBL; AJ277248; CAC19409.1; -
 CC EMBL; AF279437; AAG22064.1; -
 CC EMBL; AF387519; AAK52468.1; -
 CC EMBL; AY358890; AAG89249.1; -
 CC Genew; HGNC:14900; IL22.
 CC MW; 605330; -
 CC GO; GO:0005576; C:extracellular; IC.
 CC GO; GO:0045518; F:interleukin-22 receptor binding; NAS.
 CC GO; GO:0006953; P:acute-phase response; NAS.
 CC GO; GO:0007267; P:cell-cell signaling; IC.
 CC GO; GO:0006954; P:inflammatory response; NAS.
 CC InterPro; IPR000098; Interleukin 10.
 CC PROSITE; PS00520; INTERLEUKIN 10; 1.
 KW Cytokine; Glycoprotein; Signal; Polymorphism.
 FT SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 179 INTERLEUKIN-22.
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 68 68 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 158 158 S -> G.
 FT FTID=VAR_013078.
 SQ SEQUENCE 179 AA; 20011 MW; 3C35E64D60CF8767 CRC64;
 Query Match 0.7%; Score 8; DB 1; Length 179;
 Best Local Similarity 100.0%; Pred. No. 9.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

QY      1153 LLLLLLV 1160
Db      21 LLLLLLV 28

RESULT 18
EMB2_CAVPO
ID      EMB2_CAVPO      STANDARD;      PRT;      234 AA.
AC      P25709;
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Eosinophil granule major basic protein 2 precursor (MBP-2).
GN      MBP2.
OS      Cavia porcellus (Guinea pig).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX      NCBI_TaxID=10141;

[1]
RN      SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP      TISSUE=Eosinophil;
RX      MEDLINE=91224343; PubMed=2026266;
RA      Aoki I., Shindoh Y., Nishida T., Nakai S., Hong Y.-M., Mio M.,
RA      Saito T., Tasaka K.;
RT      Comparison of the amino acid and nucleotide sequences between human
RT      and two guinea pig major basic proteins.";
RL      FEBS Lett. 282:56-60(1991).
CC      -1- FUNCTION: MBP may play some important roles in the allergic
CC      reactions and inflammations, since MBP is capable of releasing
CC      histamine from mast cells and damaging the epithelial cells of
CC      bronchial tubes. Antiparasitic and antibiotic.
CC      -1- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC      granule (crystalloid core).
CC      -1- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; D00817; BAA00697.1; -
CC      PIR; S15102; S15102.
CC      HSPF; P13727; I180.
CC      InterPro; IPR002352; Emaior basic.
CC      InterPro; IPR001304; Lectin_C.
CC      Pfam; PF00059; lectin_C; 1.
CC      PRINTS; PR00770; EMAJORBASICP.
CC      SMART; SM00034; CLECT; 1.
CC      PROSITE; PS00615; C-TYPE LECTIN_1; 1.
CC      PROSITE; PS00041; C-TYPE LECTIN_2; 1.
CC      Eosinophil; Signal; Immune response; Antibiotic; Lectin;
CC      Multigene family; Glycoprotein.
CC      SIGNAL 1 15 POTENTIAL.
CC      PROPEP 16 115 ACIDIC.
CC      CHAIN 116 234 EOSINOPHIL GRANULE MAJOR BASIC PROTEIN 2.
CC      DOMAIN 133 234 C-TYPE LECTIN (SHORT FORM).
CC      DISULFID 135 232 BY SIMILARITY.
CC      DISULFID 209 224 BY SIMILARITY.
CC      CARBOHYD 69 69 O-LINKED (GLYCOSAMINOGLYCAN) (BY
CC      SIMILARITY).
CC      SEQUENCE 234 AA; 26140 MW; 7D926A942BF5116F CRC64;

Query Match 0.7%; Score 8; DB 1; Length 234;
Best Local Similarity 100.0%; Pred.No.12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1153 LLLLLLV 1160
Db      4 LLLLLLV 11

RESULT 19
COMT_RAT
ID      COMT_RAT      STANDARD;      PRT;      264 AA.
AC      P22734;
DT      01-AUG-1991 (Rel. 19, Created)
DT      01-MAY-1992 (Rel. 22, Last sequence update)
DT      10-OCT-2003 (Rel. 42, Last annotation update)
DE      Catechol O-methyltransferase, membrane-bound form (EC 2.1.1.6)
DE      (MB-COMT) [Contains: Catechol O-methyltransferase, soluble form
DE      (S-COMT)].
GN      COMT.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;

[1]
RN      SEQUENCE FROM N.A.
RP      TISSUE=Liver;
RX      MEDLINE=94107221; PubMed=8280056;
RA      Tenhunen J., Ulanen I.;
RT      "Production of rat soluble and membrane-bound catechol O-
RT      methyltransferase forms from bifunctional mRNAs.";
RL      Biochem. J. 296:595-600(1993).

[2]
RN      SEQUENCE OF 11-264 FROM N.A.
RP      MEDLINE=91033034; PubMed=2227437;
RA      Salminen M., Lundstroem K., Tilgmann C., Savolainen R., Kalkkinen N.,
RA      Ulanen I.;
RT      "Molecular cloning and characterization of rat liver catechol-O-
RT      methyltransferase.";
RL      Gene 93:241-247(1990).

[3]
RN      SEQUENCE OF 1-10 FROM N.A., AND CHARACTERIZATION OF THE TWO FORMS.
RP      MEDLINE=92111472; PubMed=1765063;
RA      Ulanen I., Lundstroem K.;
RT      "Cell-free synthesis of rat and human catechol O-methyltransferase.
RT      Insertion of the membrane-bound form into microsomal membranes in
RT      vitro.";
RL      Eur. J. Biochem. 202:1013-1020(1991).

[4]
RN      X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF SOLUBLE FORM.
RP      MEDLINE=94173341; PubMed=8127373;
RA      Vidgren J., Svensson L.A., Liljas A.;
RT      "Crystal structure of catechol O-methyltransferase.";
RL      Nature 368:354-358(1994).
CC      -1- FUNCTION: Catalyzes the O-methylation, and thereby the
CC      inactivation, of catecholamine neurotransmitters and catechol
CC      hormones. Also shortens the biological half-lives of certain
CC      neuroactive drugs, like L-DOPA, alpha-methyl DOPA and
CC      isoproterenol.
CC      -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + catechol = S-
CC      adenosyl-L-homocysteine + guaiacol.
CC      -1- COFACTOR: Magnesium.
CC      -1- SUBCELLULAR LOCATION: Cytoplasmic (isoform S-COMT). Type II
CC      membrane protein (isoform MB-COMT).
CC      -1- ALTERNATIVE PRODUCTS:
CC      Event=Alternative initiation;
CC      Comment=2 isoforms, Membrane-bound/MB-COMT (shown here) and
CC      Soluble/S-COMT, are produced by alternative initiation;
CC      -1- PTM: The N-terminus is blocked.
CC      -1- SIMILARITY: TO OTHER MAMMALIAN CATECHOL-O-METHYLTRANSFERASE.
CC
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
CC      EMBL; Z12651; CAA78276.1; -.

```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales; OC

DR 28-FEB-2003 (Rel. 41, Last sequence update)
 DR 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Sphingomyelin phosphodiesterase 2 (EC 3.1.1.4.12) (Neutral
 DE sphingomyelinase) (nSMase) (Lyso-platelet activating factor-
 DE phospholipase C) (Lyso-PAF-PLC)
 GN SMPD2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND TISSUE SPECIFICITY.
 RX MEDLINE=98188255; PubMed=9520418;
 RA Tomiuk S., Hofmann K., Nix M., Zumbansen M., Stoffel W.;
 RT "Cloned mammalian neutral sphingomyelinase: functions in sphingolipid
 RT signaling?";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:3638-3643(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H.A., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.A., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Scheraga A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=20076490; PubMed=10608884;
 RA Sawai H., Domae N., Nagan N., Hannun Y.A.;
 RT "Function of the cloned putative neutral sphingomyelinase as
 RT Lyso-platelet activating factor-phospholipase C";
 RL J. Biol. Chem. 274:38131-38139(1999).
 CC -!- FUNCTION: Converts sphingomyelin to ceramide. Hydrolyze 1-acyl-2-
 CC lyso-sn-glycero-3-phosphocholine (lyso-PC) and 1-O-alkyl-2-lyso-
 CC sn-glycero-3-phosphocholine (lyso-platelet activating factor). The
 CC physiological substrate seems to be Lyso-PAF.
 CC -!- CATALYTIC ACTIVITY: Sphingomyelin + H₂O = N-acylsphingosine +
 CC choline phosphate.
 CC -!- COFACTOR: Magnesium.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- MISCELLANEOUS: This protein has an optimum pH of 6.5-7.5.
 CC -!- SIMILARITY: Belongs to the neutral sphingomyelinase family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AJ222801; CAA10995.1; -;
 CC EMBL; BC000038; AAH00038.1; -;
 CC Genbank; HGNC:11121; SMPD2.
 CC MIM; 603498; -;
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.
 CC GO; GO:0004767; F:sphingomyelin phosphodiesterase activity; TAS.

DR GO; GO:0006684; P:sphingomyelin metabolism; TAS.
 DR InterPro; IPR005135; Exo_endo_phos.
 DR Pfam; PFO372; Exo_endo_phos; 1.
 DE Hydrolase; Transmembrane; Magnesium.
 FT TRANSMEM 330 350 POTENTIAL.
 FT TRANSMEM 354 374 POTENTIAL.
 FT METAL 49 49 MAGNESIUM (BY SIMILARITY).
 FT SITE 180 180 IMPORTANT FOR SUBSTRATE RECOGNITION (BY
 FT ACT SITE 272 272 SIMILARITY).
 FT CONFLICT 3 3 GENERAL BASE (BY SIMILARITY).
 SQ SEQUENCE 423 AA; 47592 MW; 05252A923E363171 CRC64;
 L -> P (IN REF. 2).
 Query Match 0.7%; Score 8; DB 1; Length 423;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1152 GLLLLALL 1159
 DB 334 GLLLLALL 341
 RESULT 22
 ENGA_MYCPE STANDARD; PRT; 444 AA.
 ID ENGA_MYCPE
 AC Q8EWH6;
 DT 15-MAR-2004 (Rel. 43, Created)
 DT 15-MAR-2004 (Rel. 43, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE GTP-binding protein engA.
 DE ENGA OR MYPE2290.
 GN Mycoplasma penetrans.
 OS Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.
 OX NCBI_TaxID=28227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HF-2;
 RX MEDLINE=22354719; PubMed=12466555;
 RA Sasaki Y., Ishikawa J., Yamashita A., Oshima K., Kenri T., Furuya K.,
 RA Yoshino C., Horino A., Shiba T., Sasaki T., Hattori M.;
 RT "The complete genomic sequence of Mycoplasma penetrans, an
 RT intracellular bacterial pathogen in humans.";
 RL Nucleic Acids Res. 30:5293-5300(2002).
 CC -!- FUNCTION: GTPase of unknown physiological role.
 CC -!- SIMILARITY: Belongs to the era/trmE family of GTP-binding
 CC proteins. EngA subfamily.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AP004170; BAC44020.1; -;
 CC HAMAP; MF_00195; -; 1.
 DR InterPro; IPR005289; GTP-binding_dom.
 DR InterPro; IPR002917; MMR_HSR1.
 DR InterPro; IPR001806; Ras_trnsmfng.
 DR InterPro; IPR005225; Small_GTP.
 DR Pfam; PF01926; MMR_HSR1; 1.
 DR PRINTS; PR00449; RASTRNSFRNG.
 DR TIGRFAMs; TIGR00650; MG442; 2.
 DR TIGRFAMs; TIGR00231; small_GTP; 2.
 DR GTP-binding; Repeat; Complete proteome.
 KW NP_BIND 8 15 GTP 1 (POTENTIAL).
 FT NP_BIND 55 59 GTP 1 (POTENTIAL).
 FT NP_BIND 118 121 GTP 1 (POTENTIAL).
 FT NP_BIND 179 186 GTP 2 (POTENTIAL).
 FT NP_BIND 226 230 GTP 2 (POTENTIAL).
 FT NP_BIND 291 294 GTP 2 (POTENTIAL).
 SQ SEQUENCE 444 AA; 51231 MW; A7B636F43D7CB1BE CRC64;

```

Query Match          0.7%; Score 8; DB 1; Length 444;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 165 IIVLDGS 172
DB 258 IIVLDGS 265

RESULT 23
HISX COREP
ID HISX COREP STANDARD; PRT; 451 AA.
AC Q8FNZ0;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hispidinol dehydrogenase (SC 1.1.1.23) (HDH).
GN HISD OR CE2003.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RX MEDLINE=22723752; PubMed=12840036;
RA Nishio Y., Nakamura Y., Kawarabayashi Y., Usuda Y., Kimura E.,
RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ireo K.,
RA Gojcbori T.;
RT "Comparative complete genome sequence analysis of the amino acid
RT replacements responsible for the thermostability of Corynebacterium
RT efficiens.";
RL Genome Res. 13:1572-1579(2003).
CC -!- FUNCTION: Catalyzes the sequential NAD-dependent oxidations of L-
CC histidinol to L-histidinaldehyde and then to L-histidine (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: L-histidinol + 2 NAD(+) + H(2)O = L-histidine
CC + 2 NADH.
CC -!- COFACTOR: Binds 1 zinc ion per subunit (By similarity).
CC -!- PATHWAY: Histidine biosynthesis; ninth (last) step.
CC -!- SIMILARITY: Belongs to the hispidinol dehydrogenase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC ENBL; AP005220; BAC18813.1; -.
CC HAMAP; MF 01024; -.
CC InterPro; IPR001692; Histidinol dh.
CC Pfam; PF00815; Histidinol dh; 1.
CC PRINTS; PR00083; HOLDHGRNASE.
CC ProDom; PD002680; Histidinol dh.
CC TIGRfams; TIGR00069; hisd; 1.
CC PROSITE; PS00611; HISOL_DEHYDROGENASE; FALSE NEG.
KW Histidine biosynthesis; Oxidoreductase; NAD; Metal-binding; Zinc;
KW Complete proteome.
FT ACT_SITE 332 BY SIMILARITY.
FT ACT_SITE 333 BY SIMILARITY.
FT METAL 263 ZINC (BY SIMILARITY).
FT METAL 266 ZINC (BY SIMILARITY).
FT METAL 366 ZINC (BY SIMILARITY).
FT METAL 425 ZINC (BY SIMILARITY).
SQ SEQUENCE 451 AA; 47770 MW; 5A2D5AA7B990ECCF CRC64;

Query Match          0.7%; Score 8; DB 1; Length 451;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 334 DEAAALKDI 341
DB 402 DEAAALKDI 409

RESULT 24
D4DR_HUMAN
ID D4DR_HUMAN STANDARD; PRT; 467 AA.
AC P21917;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE D(4) dopamine receptor (D(2C) dopamine receptor).
DR D4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ALLELE D4.7).
RX MEDLINE=92310588; PubMed=1319557;
RA van Tol H.H., Wu C.M., Guan H.-C., Ohara K., Bunzow J.R.,
RA Civelli O., Kennedy J., Seeman P., Niznik H.B., Jovanovic V.;
RT "Multiple dopamine D4 receptor variants in the human population.";
RL Nature 358:149-152(1992).
RN [2]
RP SEQUENCE FROM N.A. (ALLELE D4.2).
RC TISSUE=Brain;
RX MEDLINE=91204054; PubMed=1840645;
RA van Tol H.H.M., Bunzow J.R., Guan H.-C., Sunahara R.K., Seeman P.,
RA Niznik H.B., Civelli O.;
RT "Cloning of the gene for a human dopamine D4 receptor with high
RT affinity for the antipsychotic clozapine.";
RL Nature 350:610-614(1991).
RN [3]
RP 3D-STRUCTURE MODELING.
RX MEDLINE=93038566; PubMed=1359063;
RA Livingston C.D., Strange P.G., Naylor L.H.;
RT "Molecular modelling of D2-like dopamine receptors.";
RL Biochem. J. 287:277-282(1992).
RN [4]
RP VARIANT GLY-194.
RX MEDLINE=95243275; PubMed=7726213;
RA Seeman P., Ulpian C., Chouinard G., van Tol H.H.M., Dwosh H.,
RA Lieberman J.A., Sininovich X., Liu I.S.C., Wayne J., Voruganti P.,
RA Hudson C., Serjeant G.R., Masibay A.S., Seeman M.V.;
RT "Dopamine D4 receptor variant, D4-glycine-194, in Africans, but not in
RT Caucasians: no association with schizophrenia.";
RL Am. J. Med. Genet. 54:384-390(1994).
CC -!- FUNCTION: This is one of the five types (D1 to D5) of receptors
CC for dopamine. The activity of this receptor is mediated by G
CC proteins which inhibit adenylyl cyclase.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- POLYMORPHISM: The number of repeats is highly polymorphic and
CC varies among different alleles. The sequence shown is that of
CC allele D4.7.
CC
CC -!- SIMILARITY: Belongs to family 1 of G-protein coupled receptors.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC ENBL; L12398; AAB59386.1; -.
CC ENBL; L12397; -. NOT_ANNOTATED_CDS.
CC PIR; S15079; DTHUD4.
CC Genew; HGNC:3025; DRD4.
CC MIM; 126452; -.
CC GO; GO:0005887; C:integral to plasma membrane; TAS.
CC GO; GO:0004952; F:dopamine receptor activity; TAS.

```



```

Query Match          0.7%; Score 8; DB 1; Length 487;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      626 ALGNAVIL 633
      |||||
DB      171 ALGNAVIL 178

RESULT 28
CPBI RAT
ID CPBI RAT STANDARD; PRT; 491 AA.
AC P00176;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last annotation update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B1 (EC 1.14.14.1) (CYP1B1) (P450-B) (P450-PB1 and
DE P450-PB2) (P450-LM2).
GN CYP2B1 OR CYP2B-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OC NCBI_TaxID=10116;
RN [1]
RN SEQUENCE OF 6491 FROM N.A. (ISOZYMES PB1 AND PB2).
RP MEDLINE=6222224; PubMed=6953431;
RA Fujii-Kuriyama Y., Mizukami Y., Kawajiri K., Sogawa K., Muramatsu M.;
RT "Primary structure of a cytochrome P-450: coding nucleotide sequence
RT of phenobarbital-inducible cytochrome P-450 cDNA from rat liver.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:2793-2797(1982).
RN [2]
RN REVISIONS TO 166; 292 AND 378 (ISOZYMES PB1 AND PB2).
RA Fujii-Kuriyama Y., Mizukami Y., Kawajiri K., Sogawa K., Muramatsu M.;
RL Proc. Natl. Acad. Sci. U.S.A. 79:5443-5443(1982).
RN [3]
RN SEQUENCE OF 1-22.
RP MEDLINE=79194111; PubMed=109438;
RA Botelho L.H., Ryan D.E., Levin W.;
RT "Amino acid compositions and partial amino acid sequences of three
RT highly purified forms of liver microsomal cytochrome P-450 from rats
RT treated with polychlorinated biphenyls, phenobarbital, or 3-
RT methylcholanthrene.";
RL J. Biol. Chem. 254:5635-5640(1979).
RN [4]
RP PHOSPHORYLATION
RA Peyerlin W., Taniguchi H.;
RT "Phosphorylation of hepatic phenobarbital-inducible cytochrome
RT P-450.";
RL EMBO J. 8:3003-3010(1989).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- INDUCTION: By phenobarbital.
CC -!- PTM: Phosphorylation is accompanied by a decrease in enzyme
CC activity.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announcement
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: J00719; AAA41024.1; -.

```

DR EMBL; M37134; ABC42028.1; -
 DR PIR; A00176; O4RTP2.
 DR PIR; A21162; O4RTP2.
 DR HSP; P00179; 1DT6.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008068; EP450_CYP2B.
 DR Pfam; PF00067; P450; 1.
 DR PRINTS; PR01685; EP450CYP2B.
 DR PRINTS; PR00385; P450.
 DR PROSITE; PS00086; CYTOCHROME_P450; 1.
 KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
 KW Microsome; Endoplasmic reticulum; Phosphorylation.
 FT MOD_RES 128 128 PHOSPHORYLATION (BY PKA).
 FT METAL 436 436 IRON (HEME AXIAL LIGAND).
 FT VARIANT 303 303 S -> G (IN ISOZYME PB2).
 FT VARIANT 321 322 AE -> TV (IN ISOZYME PB2).
 FT VARIANT 337 337 L -> P (IN ISOZYME PB2).
 FT VARIANT 339 339 T -> S (IN ISOZYME PB2).
 FT VARIANT 344 344 S -> T (IN ISOZYME PB2).
 SQ SEQUENCE 491 AA; 55933 MW; 74615501AD5497DD CRC64;
 Query Match 0.78; Score 8; DB 1; Length 491;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1153 LLLALLLV 1160
 Db 6 LLLALLLV 13
 RESULT 29
 CYP2B RAT STANDARD; PRT; 491 AA.
 AC P04157; O64582;
 DT 20-MAR-1987 (Rel. 04, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Cytochrome P450 2B2 (BC 1.14.14.1) (CYP11B2) (P450E) (P450 PB4).
 GN CYP2B2 OR CYP2B-2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 SEQUENCE FROM N.A.
 RP MEDLINE=83247397; PubMed=6306654;
 RA Mizukami Y., Sogawa K., Suwa Y., Muramatsu M., Fujii-Kuriyama Y.;
 RT "Gene structure of a phenobarbital-inducible cytochrome P-450 in rat liver."
 RL Proc. Natl. Acad. Sci. U.S.A. 80:3958-3962(1983).
 RN [2]
 SEQUENCE.
 RP MEDLINE=86059379; PubMed=3877725;
 RA Frey A.B., Waxman D.J., Kreibich G.;
 RT "The structure of phenobarbital-inducible rat liver cytochrome P-450 isoenzyme PB-4. Production and characterization of site-specific antibodies."
 RL J. Biol. Chem. 260:15253-15265(1985).
 RN [3]
 SEQUENCE OF 168-491 FROM N.A.
 RP MEDLINE=84159487; PubMed=6689485;
 RA Phillips I.R., Shephard E.A., Ashworth A., Rabin B.R.;
 RT "Cloning and sequence analysis of a rat liver cDNA coding for a phenobarbital-inducible microheterogeneous cytochrome P-450 variant: regulation of its messenger level by xenobiotics."
 RL Gene 26:41-52(1983).
 RN [4]
 SEQUENCE OF 281-491 FROM N.A.
 RP MEDLINE=83291091; PubMed=6688421;
 RA Kumar A., Raphael C., Adesnik M.;
 RT "Cloned cytochrome P-450 cDNA. Nucleotide sequence and homology to multiple phenobarbital-induced mRNA species."
 RL J. Biol. Chem. 258:11280-11284(1983).

DR EMBL; J00728; AAA41056.1; -
 DR EMBL; J00720; AAA41056.1; JOINED.
 DR EMBL; J00721; AAA41056.1; JOINED.
 DR EMBL; J00722; AAA41056.1; JOINED.
 DR EMBL; J00723; AAA41056.1; JOINED.
 DR EMBL; J00724; AAA41056.1; JOINED.
 DR EMBL; J00725; AAA41056.1; JOINED.
 DR EMBL; J00726; AAA41056.1; JOINED.
 DR EMBL; K00596; AAA41029.1; -
 DR EMBL; K01626; AAA41037.1; -
 DR EMBL; K01721; AAA41026.1; -
 DR EMBL; D00250; BAA00181.1; -
 DR EMBL; M13234; AAA41057.1; -
 DR PIR; A21162; O4RTP2.
 DR PIR; A60822; A60822.
 DR HSP; P00179; 1DT6.
 DR InterPro; IPR001128; Cytochrome P450.
 DR InterPro; IPR008068; EP450_CYP2B.
 [5]
 RP ERRATUM.
 RA Kumar A., Raphael C., Adesnik M.;
 RL J. Biol. Chem. 259:6039-6039(1984).
 [6]
 RP SEQUENCE OF 323-431 FROM N.A.
 RX MEDLINE=86205943; PubMed=3458196;
 RA Atchison M.L., Adesnik M.;
 RT "Gene conversion in a cytochrome P-450 gene family."
 RL Proc. Natl. Acad. Sci. U.S.A. 83:2300-2304(1986).
 [7]
 RP SEQUENCE OF 385-491 FROM N.A.
 RX MEDLINE=84153837; PubMed=6322758;
 RA Affolter M., Anderson A.;
 RT "Segmental homologies in the coding and 3' non-coding sequences of rat liver cytochrome P-450e and P-450b cDNAs and cytochrome P-450e-like genes."
 RL Biochem. Biophys. Res. Commun. 118:655-662(1984).
 [8]
 RP SEQUENCE OF 1-20 FROM N.A.
 RX MEDLINE=88273074; PubMed=2839467;
 RA Hashimoto T., Matsumoto T., Nishizawa M., Kawabata S., Morohashi K., Handa S., Omura T.;
 RT "A mutant rat strain deficient in induction of a phenobarbital-inducible form of cytochrome P-450 in liver microsomes."
 RL J. Biochem. 103:487-492(1988).
 [9]
 RP PHOSPHORYLATION.
 RX MEDLINE=90059885; PubMed=2583091;
 RA Pyerin W., Taniguchi H.;
 RT "Phosphorylation of hepatic phenobarbital-inducible cytochrome P-450."
 RL EMBO J. 8:3003-3010(1989).
 CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate monooxygenases. In liver microsomes, this enzyme is involved in an NADPH-dependent electron transport pathway. It oxidizes a variety of structurally unrelated compounds, including steroids, fatty acids, and xenobiotics.
 CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH + oxidized flavoprotein + H(2)O.
 CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
 CC -!- INDUCTION: By phenobarbital.
 CC -!- PTM: Phosphorylation is accompanied by a decrease in enzyme activity.
 CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).


```
DR Pfam; PF00067; P450; 1.
DR PRINTS; PR01685; EP450ICYP2B.
DR PROSITE; PS00086; CYTOCHROME P450.
DR PROSITE; PS00086; CYTOCHROME P450; FALSE NEG.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation.
FT MOD_RES 128 128 PHOSPHORYLATION (BY PKA).
FT METAL 436 436 IRON (HEME AXIAL LIGAND).
FT CONFLICT 292 292 L -> P (IN REF. 2).
FT CONFLICT 321 321 T -> A (IN REF. 2 AND 4).
FT CONFLICT 322 322 E -> V (IN REF. 1).
FT CONFLICT 438 438 G -> D (IN REF. 4).
FT CONFLICT 444 444 N -> K (IN REF. 3).
FT CONFLICT 473 473 K -> M (IN REF. 1).
FT CONFLICT 476 476 G -> D (IN REF. 2 AND 4).
SQ SEQUENCE 491 AA; 55932 MW; 00CB6937FDD44BC CRC64;

Query Match 0.7%; Score 8; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLIV 1160
D 6 LLLALLIV 13

RESULT 30
CPBA_MOUSE STANDARD; PRT; 500 AA.
AC P12791;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Cytochrome P450 2B10 (EC 1.14.14.1) (CYP1B10) (Testosterone 16-alpha
DE hydroxylase) (P450-16-alpha) (Clone PF3/46).
GN CYP2B10 OR CYP2B-10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=69118235; PubMed=3219345;
RX Neshiro M., Lakso M., Kawajiri K., Negishi M.;
RT "Rip locus: regulation of female-specific isozyme (I-P-450(16 alpha)
RT of testosterone 16 alpha-hydroxylase in mouse liver, chromosome
RT localization, and cloning of p-450 CDNA.";
RL Biochemistry 27:6434-6443 (1988).
CC -!- FUNCTION: Cytochromes P450 are a group of heme-thiolate
CC monooxygenases. In liver microsomes, this enzyme is involved in an
CC NADPH-dependent electron transport pathway. It oxidizes a variety
CC of structurally unrelated compounds, including steroids, fatty
CC acids, and xenobiotics.
CC -!- CATALYTIC ACTIVITY: RH + reduced flavoprotein + O(2) = ROH +
CC oxidized flavoprotein + H(2)O.
CC -!- SUBCELLULAR LOCATION: Membrane-bound. Endoplasmic reticulum.
CC -!- SIMILARITY: Belongs to the cytochrome P450 family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M21856; AAA0425.1; -.
DR PIR; B31047; B31047.
DR HSSP; P00179; 1DT6.
DR MGD; MGI:88598; Cyp2b10.
DR InterPro; IPR001128; Cytochrome P450.
DR Pfam; PF00067; P450; 1.
DR
```

```
DR PRINTS; PR01685; EP450ICYP2B.
DR PRINTS; PR00385; P450.
DR PROSITE; PS00086; CYTOCHROME P450; FALSE NEG.
KW Oxidoreductase; Monooxygenase; Electron transport; Membrane; Heme;
KW Microsome; Endoplasmic reticulum; Phosphorylation.
FT MOD_RES 128 128 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
FT METAL 445 445 IRON (HEME AXIAL LIGAND).
SQ SEQUENCE 500 AA; 56743 MW; F660A008D0FBA94 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 500;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLIV 1160
D 6 LLLALLIV 13

RESULT 31
VE1_BP2V STANDARD; PRT; 604 AA.
AC P11298;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Replication protein E1.
DE Replication protein E1.
GN E1.
OS Bovine papillomavirus type 2.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10560;
RN [1]
RP SEQUENCE FROM N.A.
RA Groff D.E., Mitra R., Lancaster W.D.;
RL Submitted (MAY-1989) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M20219; AAA66833.1; -.
DR PIR; C31169; W1WLB2.
DR InterPro; IPR001177; Papillom_E1.
DR Pfam; PF00519; E1; 1.
DR Pfam; PF00524; E1_N; 1.
KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
KW Nuclear protein.
FT NP_BIND 432 439 ATP (POTENTIAL).
SQ SEQUENCE 604 AA; 68077 MW; D2D7036ADE88A9DD CRC64;

Query Match 0.7%; Score 8; DB 1; Length 604;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 927 LAAGSDSN 934
D 336 LAAGSDSN 343

RESULT 32
VE1_BP2V STANDARD; PRT; 605 AA.
ID VE1_BP2V
AC P03116; Q9WVH1;
DT 21-JUL-1986 (Rel. 01, Created)
DR
```

DT 21-JUL-1996 (Rel. 01, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 GN Replication protein E1.
 GN E1.
 OS Bovine papillomavirus type 1.
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
 OC Papillomavirus.
 OX NCBI_TaxID=10559;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=63012974; PubMed=6289124;
 RA Chen E.Y., Howley P.M., Levinson A.D., Seeburg P.H.;
 RT "The primary structure and genetic organization of the bovine
 RT papillomavirus type 1 genome.";
 RL Nature 299:529-534(1982).
 RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91122053; PubMed=1846806;
 RA Ustav M., Stenlung A.;
 RT "transient replication of BPV-1 requires two viral polypeptides
 RT encoded by the E1 and E2 open reading frames.";
 RL EMBO J. 10:449-457(1991).
 RN (3)
 RP CHARACTERIZATION.
 RX MEDLINE=93281701; PubMed=9389467;
 RA Yang L., Mohr I., Fouts E., Lim D.A., Nohale M., Botchan M.;
 RT "The E1 protein of bovine papilloma virus 1 is an ATP-dependent DNA
 RT helicase.";
 RL Proc. Natl. Acad. Sci. U.S.A. 90:5086-5090(1993).
 CC -!- FUNCTION: ATP-DEPENDENT DNA HELICASE REQUIRED FOR INITIATION OF
 CC VIRAL DNA REPLICATION. IT FORMS A COMPLEX WITH THE VIRAL E2
 CC PROTEIN. THE E1-E2 COMPLEX BINDS TO THE REPLICATION ORIGIN WHICH
 CC CONTAINS BINDING SITES FOR BOTH PROTEINS.
 CC -!- SUBCELLULAR LOCATION: Nuclear.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X02346; CAB46511.1; -;
 DR PIR; A03663; WIMLEB.
 DR PDB; 1F08; 16-WAY-01.
 DR TRANSFAC; T0203; -;
 DR InterPro; IPR001177; Papillom_E1.
 DR Pfam; PF00519; E1; 1.
 DR Pfam; PF00524; E1 N; 1.
 KW Early protein; DNA replication; Helicase; ATP-binding; DNA-binding;
 KW Nuclear protein; 3D-structure.
 FT NP BIND 433 440 ATP (POTENTIAL)
 FT SEQUENCE 605 AA; 65190 MW; C8400B758F60650B CRC64;
 SQ
 Query Match 0.7%; Score 8; DB 1; Length 605;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 927 LAAGSDSN 934
 Db 337 LAAGSDSN 344
 |||||
 RESULT 33
 TMS6_HUMAN STANDARD; PRT; 811 AA.
 ID TMS6_HUMAN
 AC Q8IU80; Q8IU82; Q8IXV8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Transmembrane protease, serine 6 (BC 3.4.21.-) (Matrptase-2).
 GN TMPRSS6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
 RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Levenson M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., McCann O.T.,
 RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehara H.K., Skuce C.D., Smalley S., Smith M.I.,
 RA Soderlund C., Spragon I., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Dorman A., Fang P., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Mala J.E., Nguyen T., Pan H.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Z., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
 RA Korf I., Bedell J.A., Hillier L.W., Marais E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
 RA Budarf M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edelman L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
 RA Peyrard M., Kedar D., Serousai E., Fransson I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tlahun Y., Wright H.;
 RT "The DNA sequence of human chromosome 22.";
 RL Nature 402:489-495(1999).
 RN (3)
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX TISSUE=Brain;
 MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Vallalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

FT				(POTENTIAL).
FT	DOMAIN	77	811	EXTRACELLULAR (POTENTIAL).
FT	DOMAIN	213	336	CUB 1.
FT	DOMAIN	335	452	CUB 2.
FT	DOMAIN	457	489	LDL-RECEPTOR CLASS A 1.
FT	DOMAIN	490	526	LDL-RECEPTOR CLASS A 2.
FT	DOMAIN	530	567	LDL-RECEPTOR CLASS A 3.
FT	DOMAIN	577	811	SERINE PROTEASE.
FT	ACT_SITE	617	617	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	668	668	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	ACT_SITE	762	762	CHARGE RELAY SYSTEM (BY SIMILARITY).
FT	CARBOHYD	136	136	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	184	184	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	216	216	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	338	338	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	433	433	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	453	453	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD	518	518	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	VARSPLIC	409	461	LCGLRLQLQYAKRIPVNAVTAAGITINFGSISLTGPGVRVHY GLYNQSDPCPE -> YHFLSLMLPLFPPPSLPSSVTVP SLAEQVNLRGAARGSGWGWCQACC (in isoform 2).
FT	VARSPLIC	462	811	/FTId=VSP_008379. Missing [in isoform 2]. /FTId=VSP_008380. A -> V [IN REF. 4].
FT	CONFLICT	116	116	A -> V [IN REF. 4].
FT	SEQUENCE	811 AA;	89999 MW;	7EEF193F65DD9D CRC64;
Query Match		0.7%;	Score 8;	DB 1; Length 811;
Best Local Similarity		100.0%;	Pred. No. 33;	
Matches		8; Conservative	0; Mismatches	0; Indels
				0; Gaps
QY		1155 LLALLVLA 1162		
Db		59 LLALLVLA 66		
RESULT 34				
ID	ITAS_XENLA	STANDARD;	PRT;	1050 AA.
ID	ITAS_XENLA	AC Q06274;		
DT	15-JUL-1998	(Rel. 36, Created)		
DT	15-JUL-1998	(Rel. 36, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Integrin alpha-5 precursor (fibronectin receptor alpha subunit)			
DE	(integrin alpha-F) (VLA-5).			
OS	Xenopus laevis (African clawed frog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;			
OC	Xenopodinae; Xenopus.			
OX	NCBI_TaxID=8355;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RP	MEDLINE=9534994; PubMed=7619730;			
RA	Jos T.O., Whittaker C.A., Meng F., Desimone D.W., Gnau V.,			
RA	Whittaker P.;			
RT	"Integrin alpha 5 during early development of Xenopus laevis.";			
RL	Mech. Dev. 50:187-199(1995).			
RL	[2]			
RP	SEQUENCE OF 318-393 FROM N.A.			
RX	MEDLINE=94008528; PubMed=8404528;			
RA	Whittaker C.A., Desimone D.W.;			
RT	"Integrin alpha subunit mRNAs are differentially expressed in early			
RL	Xenopus embryos.";			
RL	Development 117:1239-1249(1993).			
CC	-I- FUNCTION: INTEGRIN ALPHA-5/BETA-1 IS A RECEPTOR FOR FIBRONECTIN.			
CC	-I- IT RECOGNIZES THE SEQUENCE R-G-D IN ITS LIGANDS.			
CC	-I- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA			
CC	SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A			
CC	DISULFIDE BOND. ALPHA-5 ASSOCIATES WITH BETA-1.			
CC	-I- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-I- SIMILARITY: Belongs to the integrin alpha chain family.			
CC	-I- SIMILARITY: Contains 7 FG-GAP repeats.			

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

 EMBL: U16883; AAA9668.1; -;
 EMBL: L10191; AAA16249.1; -;
 PIR: I51527; I51527.
 HSSP: P06756; IJUV2.
 InterPro: IPR000413; Integrin_alpha.
 Pfam: PF01839; FG-GAP; 4.
 Pfam: PF00357; integrin_A; 1.
 SMART: SM00191; int_alpha; 5.
 PROSITE: PS00242; INTEGRIN ALPHA; 1.
 Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
 Signal; Repeat.
 SIGNAL 1 32 POTENTIAL.
 CHAIN 33 1050 INTEGRIN ALPHA-5.
 CHAIN 33 932 INTEGRIN ALPHA-5 HEAVY CHAIN (POTENTIAL).
 CHAIN 933 1050 INTEGRIN ALPHA-5 LIGHT CHAIN (POTENTIAL).
 DOMAIN 33 996 EXTRACELLULAR (POTENTIAL).
 TRANSMEM 997 1022 POTENTIAL.
 DOMAIN 1023 1050 CYTOPLASMIC (POTENTIAL).
 REPEAT 48 110 FG-GAP 1.
 REPEAT 120 188 FG-GAP 2.
 REPEAT 189 246 FG-GAP 3.
 REPEAT 259 312 FG-GAP 4.
 REPEAT 313 373 FG-GAP 5.
 REPEAT 379 438 FG-GAP 6.
 REPEAT 442 494 FG-GAP 7.
 CA_BIND 324 332 POTENTIAL.
 CA_BIND 390 398 POTENTIAL.
 CA_BIND 454 462 POTENTIAL.
 SITE 1025 1029 GFFKR MOTIF.
 DISULFID 90 99 BY SIMILARITY.
 DISULFID 145 166 BY SIMILARITY.
 DISULFID 182 195 BY SIMILARITY.
 DISULFID 502 513 BY SIMILARITY.
 DISULFID 519 573 BY SIMILARITY.
 DISULFID 636 642 BY SIMILARITY.
 DISULFID 708 721 BY SIMILARITY.
 DISULFID 862 910 INTERCHAIN (BY SIMILARITY).
 CARBOHYD 917 922 BY SIMILARITY.
 CARBOHYD 95 95 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 98 98 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 287 297 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 297 297 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 306 306 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 507 507 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 515 515 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 521 521 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 600 600 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 649 649 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 714 714 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 763 763 N-LINKED (GLCNAC. .) (POTENTIAL).
 CARBOHYD 861 861 N-LINKED (GLCNAC. .) (POTENTIAL).
 SEQUENCE 1050 AA; 115961 MW; 10ED9615358D918 CRC64;

Query Match 0.7%; Score 8; DB 1; Length 1050;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 RFGSSIAS 548
 |||||
 Db 380 RFGSSIAS 387

RESULT 35
 ITAL DROME
 ID ITAL DROME STANDARD; PRT: 1146 AA.
 AC Q24247; QSVYF6;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Integrin alpha-PS1 precursor (position-specific antigen 1, alpha chain) (Protein multiple edematous wings).
 GN MEW OR CGI771.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 CX NCBI_taxid=7227;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC STRAIN=Oregon-R;
 RX MEDLINE=94059764; PubMed=8240969;
 RA Wehrli M., Diantonio A., Pearnley I.M., Smith R.J., Wilcox M.;
 RT "Cloning and characterization of alpha PSI, a novel Drosophila melanogaster integrin.";
 RL Mech. Dev. 43:21-36(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Berkley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Bencos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hootin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kiyara C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 CC -1- SUBUNIT: INTEGRIN ALPHA-PS1/BETA-PS IS A RECEPTOR FOR LAMININ.
 CC -1- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. THE ALPHA
 CC SUBUNIT IS COMPOSED OF AN HEAVY AND A LIGHT CHAIN LINKED BY A
 CC DISULFIDE BOND. ALPHA-PS1 ASSOCIATES WITH BETA-PS.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: Belongs to the integrin alpha chain family.
 CC -1- SIMILARITY: Contains 7 FG-GAP repeats.
 CC -1- CAUTION: Ref.2 sequence differs from that shown due to erroneous

CC
CC gene model prediction.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL curation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.ebi.ac.uk/termsandconditions>
CC or send an email to license@ebi.ac.uk).

[illegible]

```
Query Match      0.7%; Score 8; DB 1; Length 1146;
Best Local Similarity 100.0%; Pred.No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	1151 GGLLLLLL 1158
Db	1094 GGLLLLLL 1101

RESULT	36
ITEM	HUMAN
ID	ITAM HUMAN
IC	P11215
STANDARD;	PRT; 1152 AA.
DT	01-JUL-1989 (Rel. 11, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	10-OCT-2003 (Rel. 42, Last annotation update)
DE	Integrin alpha-M precursor (Cell surface glycoprotein MAC-1 alpha

subunit) (CR-3 alpha chain) (CD11b) (Leukocyte adhesion receptor Mol1)
DE (Neutrophil adherence receptor).
DE ITGAM OR CR3A OR CD11B.
DE
OS Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
MEDLINE=88315033; PubMed=2457584;
CORBI A.L., Kishimoto T.K., Miller L.J., Springer T.A.;
ERR "the human leukocyte adhesion glycoprotein Mac-1 (complement receptor
type 3, CD11b) alpha subunit. Cloning, primary structure, and
relation to the integrins, von Willebrand factor and factor B.";
J. Biol. Chem. 263:12403-12411(1988).
[2]
SEQUENCE FROM N.A.
MEDLINE=88190151; PubMed=2833753;
ARNAOUT M.A., Remold-O'Donnell E., Pierce M.W., Harris P., Tenen D.G.;
ERR "Molecular cloning of the alpha subunit of human and guinea pig
leukocyte adhesion glycoprotein Mol1: chromosomal localization and
homology to the alpha subunits of integrins.";
Proc. Natl. Acad. Sci. U.S.A. 85:2776-2780(1988).
[3]
SEQUENCE FROM N.A.
MEDLINE=8857215; PubMed=2454931;
ARNAOUT M.A., Gupta S.K., Pierce M.W., Tenen D.G.;
ERR "Amino acid sequence of the alpha subunit of human leukocyte adhesion
receptor Mol1 (complement receptor type 3).";
J. Cell Biol. 106:2153-2158(1988).
[4]
SEQUENCE FROM N.A.
MEDLINE=93123748; PubMed=8419480;
FLEMING J.C., Pahl H.L., Gonzalez D.A., Smith T.F., Tenen D.G.;
ERR "Structural analysis of the CD11b gene and phylogenetic analysis of
the alpha-integrin gene family demonstrate remarkable conservation of
genomic organization and suggest early diversification during
evolution.";
J. Immunol. 150:480-490(1993).
[5]
SEQUENCE OF 9-1153 FROM N.A.
MEDLINE=89098893; PubMed=2563162;
HICKSTEIN D.D., Hickey M.J., Ozols J., Baker D.M., Back A.L.,
ROTH G.J.;
ERR "cDNA sequence for the alpha M subunit of the human neutrophil
adherence receptor indicates homology to integrin alpha subunits.";
Proc. Natl. Acad. Sci. U.S.A. 86:257-261(1989).
[6]
SEQUENCE OF 1-9 FROM N.A.
MEDLINE=92073318; PubMed=1683702;
SHELLEY C.S., ARNAOUT M.A.;
ERR "The promoter of the CD11b gene directs myeloid-specific and
developmentally regulated expression.";
Proc. Natl. Acad. Sci. U.S.A. 88:10525-10529(1991).
[7]
SEQUENCE OF 1-9 FROM N.A.
TISSUE=Blood;
MEDLINE=92144986; PubMed=1346576;
PAHL H.L., ROEMERIN A.G., Tenen D.G.;
ERR "Characterization of the myeloid-specific CD11b promoter.";
Blood 79:865-870(1992).
[8]
SEQUENCE OF 17-31.
MEDLINE=87076671; PubMed=3539202;
PIERCE M.W., Remold-O'Donnell E., Todd R.F. III, ARNAOUT M.A.;
ERR "N-terminal sequence of human leukocyte glycoprotein Mol1:
conservation across species and homology to platelet IIb/IIIa.";
Biochim. Biophys. Acta 874:368-371(1986).
[9]
X-RAY CRYSTALLOGRAPHY (1.7 ANGSTROMS) OF 148-331.
MEDLINE=95717458; PubMed=7867070;
LEES J.O., RIEU P., ARNAOUT M.A., Liddington R.;
ERR "Crystal structure of the A domain from the alpha subunit of integrin
alpha M.";

RC TISSUE=Brain, and Duodenum;
RA MEDLINE=22388257; PubMed=12477932;
RA Srausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins S.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Besak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting J., Helton E., Kettelman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.F., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 178-1157 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=20277482; PubMed=10819331;
RA Nagase T., Kikuno R., Ishikawa K.-I., Hirosewa M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XVII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro."
RL DNA Res. 7:143-150(2000).
CC -!- SIMILARITY: Belongs to the DEAD box helicase family. DEAD
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; BC002575; AA02575.1; ALT_INIT.
DR EMBL; BC037964; AA037964.1; -
DR EMBL; AB040950; BAA96041.1; -
DR Genbank; HGNC:17210; DHX37.
DR InterPro; IPR004410; DEAD.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR007502; Helicase_dom.
DR Pfam; PF04408; HA2; 1.
DR SMART; SM00271; helicase_C; 1.
DR SMART; SM00487; DEXDC; 1.
DR SMART; SM00490; HELIC; 1.
DR PROSITE; PS00690; DEAD_ATP_HELICASE; FALSE_NEG.
KW Hydrolase; Helicase; ATP-binding.
FT NP BIND 275 282 ATP (POTENTIAL).
FT SITE 372 375 DEAD BOX.
FT SITE 372 375 S -> G (IN REF. 1; AA02575).
FT CONFLICT 869 869 A -> S (IN REF. 2).
FT CONFLICT 898 898 R -> Q (IN REF. 1; AA02575).
FT CONFLICT 1081 1081 YLLAEYCEWLFQAMHPDIEKAWPTTHV -> CEFDQGGV
FT CONFLICT 1130 1157 GVDNGSLRQGLCALCTVSPGLAGSGTAAQGLFAT (IN
FT REF. 2).
FT SEQUENCE 1157 AA; 129544 MW; 49332175221B30C5 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 1157;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 380 VLLGAVGA 387
DB 857 VLLGAVGA 864

RESULT 38
ITAD_HUMAN
ID ITAD_HUMAN STANDARD; PRT; 1162 AA.
AC Q13349; Q15575; Q15576;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DT Integrin alpha-D precursor (Leukointegrin alpha D) (CD11d) (ADB2).
GN ITGAD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]_TaxID=9606;
RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RX MEDLINE=96111956; PubMed=8777714;
RA Van der Vliet M., Le Trong H., Wood C.L., Moore P.F., St John T.,
RA Staunton D.E., Gallatin W.M.;
RT "A novel leukointegrin, alpha d beta 2, binds preferentially to ICAM-
RT 3.";
RL Immunity 3:683-690(1995).
RN [2]
RP SEQUENCE OF 1-235 FROM N.A.
RX MEDLINE=20187620; PubMed=10722744;
RA Noti J.D., Johnson A.K., Dillon J.D.;
RT "Structural and functional characterization of the leukocyte integrin
RT gene Cldd. Essential role of Spl and Sp3.";
RL J. Biol. Chem. 275:8959-8969(2000).
RN [3]
RP SEQUENCE OF 457-537; 571-602; 633-666; 788-834 AND 910-1125 FROM N.A.
RX MEDLINE=96257236; PubMed=8666289;
RA Wong D.A., Davis E.M., LeBeau M., Springer T.A.;
RT "Cloning and chromosomal localization of a novel gene-encoding a human
RT beta 2-integrin alpha subunit.";
RL Gene 171:291-294(1996).
RN [4]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99059842; PubMed=9841932;
RA Grayson M.H., Van der Vliet M., Sterbinsky S.A., Michael Gallatin W.,
RA Hoffman P.A., Staunton D.E., Bochner B.S.;
RT "alpha2beta2 integrin is expressed on human eosinophils and functions
RT as an alternative ligand for vascular cell adhesion molecule 1
RT (VCAM-1).";
RL J. Exp. Med. 188:2187-2191(1998).
RN [5]
RP INTERACTION WITH VCAM1.
RX MEDLINE=99370002; PubMed=10438935;
RA Van der Vliet M., Crowe D.T., Hoekstra D., Vazeux R., Hoffman P.A.,
RA Grayson M.H., Bochner B.S., Gallatin W.M., Staunton D.E.;
RT "The leukocyte integrin alpha D beta 2 binds VCAM-1: evidence for a
RT binding interface between I domain and VCAM-1.";
RL J. Immunol. 163:1984-1990(1999).
CC -!- FUNCTION: INTEGRIN ALPHA-D/BETA-2 IS A RECEPTOR FOR ICAM3 AND
CC VCAM1. MAY PLAY A ROLE IN THE ATHEROSCLEROTIC PROCESS SUCH AS
CC CLEARING LIPOPROTEINS FROM PLAQUES AND IN PHAGOCYTOSIS OF BLOOD-
CC BORNE PATHOGENS, PARTICULATE MATTER, AND SENESENT ERYTHROCYTES
CC FROM THE BLOOD.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-D
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC LINES AND SUBSETS OF PERIPHERAL BLOOD LEUKOCYTES AND STRONGLY ON
CC TISSUE-SPECIALIZED CELLS, INCLUDING MACROPHAGES FOAM CELLS WITHIN
CC ATHEROSCLEROTIC PLAQUES, AND ON SPLENIC RED PULP MACROPHAGES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; U37028; AAB38547.1; -
EMBL; U40274; AAB60634.1; -
EMBL; U40275; AAB60635.1; -
EMBL; U40276; AAB60636.1; -
EMBL; U40277; AAB60637.1; -
EMBL; U40278; AAB60638.1; -
EMBL; U40279; AAB60639.1; JOINED.
EMBL; AF187881; AAF62875.1; -
HSP; P11215; IABX.
Genew; HGNC:6146; ITCAD.
MIM; 602453; -
GO; GO:0008305; C:integrin complex; TAS.
GO; GO:0004895; P:cell adhesion receptor activity; TAS.
GO; GO:0016337; P:cell-cell adhesion; NAS.
GO; GO:0007160; P:cell-matrix adhesion; NAS.
GO; GO:0006955; P:immune response; NAS.
InterPro; IPR000413; Integrin_alpha.
InterPro; IPR002035; VWF_A.
Pfam; PF01839; FG-GAP; 3_A.
Pfam; PF00357; integrin_A; 1.
Pfam; PF00092; vwa; 1.
PRINTS; PR01185; INTEGRINA.
PRINTS; PR00453; VWFADOMAIN.
SMART; SM00191; Int_alpha; 4.
SMART; SM00327; VWA; 1.
PROSITE; PS00242; INTEGRIN_ALPHA; 1.
PROSITE; PS50234; VWFA; 1.
Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
Signal; Repeat; Calcium;
Magnesium.
SIGNAL 1 17
CHAIN 18 1162
DOMAIN 18 1100
TRANSMEM 1101 1124
DOMAIN 1125 1162
REPEAT 32 85
DOMAIN 150 332
REPEAT 350 400
REPEAT 401 452
REPEAT 454 516
REPEAT 518 576
REPEAT 581 633
CA BIND 465 473
CA BIND 530 538
CA BIND 593 601
SITE 1127 1131
DISULFID 67 74
DISULFID 106 124
DISULFID 655 710
DISULFID 769 775
DISULFID 846 861
DISULFID 994 1018
DISULFID 1023 1028
CARBOHYD 59 59
CARBOHYD 87 87
CARBOHYD 99 99
CARBOHYD 391 391
CARBOHYD 691 691
CARBOHYD 733 733
CARBOHYD 773 773
CARBOHYD 873 873
CARBOHYD 957 957
CARBOHYD 1046 1046
CONFLICT 500 500
CONFLICT 515 515
CONFLICT 825 825
L -> V (IN REF. 2).
L -> V (IN REF. 2).

FT CONFLICT 984 984 V -> A (IN REF. 2).
SQ SEQUENCE 1162 AA; 126885 MW; P296NLA3545SD77D CRC64;
Query Match 0.7%; Score 8; DB 1; Length 1162;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 56 LVVGAPLE 63
DB 47 LVVGAPLE 54
RESULT 39
ITAX_HUMAN
ID ITAX_HUMAN STANDARD; PRT; 1163 AA.
AC P20762;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Integrin alpha-X precursor (Leukocyte adhesion glycoprotein p150,95
DE alpha chain) (Leukocyte adhesion receptor p150,95) (CD11c) (Leu M5).
GN ITGAX OR CD11c.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88166645; PubMed=3327687;
RA Corbi A.L., Miller L.J., O'Connor K., Larson R.S., Springer T.A.;
RT "cDNA cloning and complete primary structure of the alpha subunit of
RT a leukocyte adhesion glycoprotein, p150,95.";
RL EMBO J. 6:4023-4028(1987).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=90153906; PubMed=2303426;
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RT "Genomic structure of an integrin alpha subunit, the leukocyte
RT p150,95 molecule.";
RL J. Biol. Chem. 265:2782-2788(1990).
RN [3]
RP ERRATUM.
RA Corbi A.L., Garcia-Aguilar J., Springer T.A.;
RX J. Biol. Chem. 265:12750-12751(1990).
RN [4]
RP SEQUENCE OF 20-43.
RX MEDLINE=87167596; PubMed=3549901;
RA Miller L.J., Wiebe M., Springer T.A.;
RT "Purification and alpha subunit N-terminal sequences of human Mac-1
RT and p150,95 leukocyte adhesion proteins.";
RL J. Immunol. 138:2381-2383(1987).
CC -!- FUNCTION: INTEGRIN ALPHA-X/BETA-2 IS A RECEPTOR FOR FIBRINOGEN. IT
CC RECOGNIZES THE SEQUENCE G-P-R IN FIBRINOGEN. IT MEDIATES CELL-CELL
CC INTERACTION DURING INFLAMMATORY RESPONSES. IT IS ESPECIALLY
CC IMPORTANT IN MONOCYTE ADHESION AND CHEMOTAXIS.
CC -!- SUBUNIT: HETERODIMER OF AN ALPHA AND A BETA SUBUNIT. ALPHA-X
CC ASSOCIATES WITH BETA-2.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN MONOCYTES AND
CC GRANULOCYTES.
CC -!- DOMAIN: THE INTEGRIN I-DOMAIN (INSERT) IS A VWFA DOMAIN. INTEGRINS
CC WITH I-DOMAINS DO NOT UNDERGO PROTEASE CLEAVAGE.
CC -!- SIMILARITY: Belongs to the integrin alpha chain family.
CC -!- SIMILARITY: Contains 1 VWFA domain.
CC -!- SIMILARITY: Contains 7 FG-GAP repeats.
CC -!- DATABASE: NAME=PROW; NOTE=CD guide CD11c entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd11c.htm".

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed.

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----

CC EMBL; M81695; AAA89180.1; -;
CC EMBL; Y00093; CAA68283.1; -;
CC EMBL; M29165; -; NOT_ANNOTATED_CDS.
CC EMBL; M29487; AAA51620.1; ALT SEQ.
CC EMBL; M29482; AAA51620.1; JOINED.
CC EMBL; M29483; AAA51620.1; JOINED.
CC EMBL; M29484; AAA51620.1; JOINED.
CC EMBL; M29485; AAA51620.1; JOINED.
CC EMBL; M29486; AAA51620.1; JOINED.
CC PIR; A36584; RWHULC.
CC PDB; 1N3Y; 18-FEB-03.
CC Genew; HGNC.6152; ITGAX.
CC MIM; 151510; -;
CC GO; GO:0008305; C: integrin complex; TAS.
CC GO; GO:0004895; F: cell adhesion receptor activity; TAS.
CC GO; GO:0004872; F: receptor activity; TAS.
CC GO; GO:0007155; P: cell adhesion; TAS.
CC GO; GO:0007397; P: histogenesis and organogenesis; TAS.
CC InterPro; IPR000413; integrin_alpha.
CC InterPro; IPR002035; VWF_A.
CC Pfam; PF01839; FG-GAP; 3.
CC Pfam; PF00357; integrin_A; 1.
CC Pfam; PF00092; vwa; 1.
CC PRINTS; PR01185; INTEGRINA.
CC PRINTS; PR00453; VWFADOMAIN.
CC SMART; SMC03191; Int_alpha; 5.
CC SMART; SMC0327; VWA; 1.
CC PROSITE; PS00242; INTEGRIN_ALPHA; 1.
CC PROSITE; PS00234; VWFA; 1.
CC InterPro; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
CC Signal; Magnesium; Calcium; Repeat; 3D-structure.
CC SIGNAL 1 19
CC CHAIN 20 1163 INTEGRIN ALPHA-X
CC DOMAIN 20 1107 EXTRACELLULAR (POTENTIAL).
CC TRANSMEM 1108 1128 POTENTIAL.
CC DOMAIN 1129 1163 CYTOPLASMIC (POTENTIAL).
CC REPEAT 34 87 FG-GAP 1.
CC REPEAT 2 2 FG-GAP 2.
CC DOMAIN 165 351 VWFA.
CC REPEAT 402 453 FG-GAP 3.
CC REPEAT 455 517 FG-GAP 4.
CC REPEAT 518 576 FG-GAP 5.
CC REPEAT 581 633 FG-GAP 6.
CC REPEAT 466 474 FG-GAP 7.
CC CA_BIND 530 538 POTENTIAL.
CC CA_BIND 593 601 POTENTIAL.
CC SITE 1131 1135 GFFKR MOTIF.
CC DISULFID 69 76 BY SIMILARITY.
CC DISULFID 108 126 BY SIMILARITY.
CC DISULFID 655 712 BY SIMILARITY.
CC DISULFID 771 777 BY SIMILARITY.
CC DISULFID 848 863 BY SIMILARITY.
CC DISULFID 998 1022 BY SIMILARITY.
CC DISULFID 1027 1032 BY SIMILARITY.
CC CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 392 392 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 697 697 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 735 735 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 899 899 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 939 939 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 1050 1050 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CONFLICT 490 490 G -> A (IN REF. 2).
CC CONFLICT 756 756 L -> D (IN REF. 2).
CC SEQUENCE 1163 AA; 127895 MW; 6C4E19CC3F62A473 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 1163;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLALLAL 1158
DB 1115 GGLLELLAL 1122
RESULT 40
TSPI_XENLA
ID TSPI_XENLA STANDARD; PRT; 1173 AA.
AC P35448;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Thrombospondin 1 precursor.
GN THBS1 OR TSPI.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Urry L.A., Ramos J., Duquette M., Desimone D.W., Lawler J.;
RL Submitted (XX-1993) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Adhesive glycoprotein that mediates cell-to-cell and
CC cell-to-matrix interactions. Can bind to fibrinogen, fibronectin,
CC laminin, type V collagen and Integrins alpha-V/beta-1, alpha-
CC V/beta-3 and alpha-IIB/beta-3 (By similarity).
CC -!- SUBUNIT: Homotrimer; disulfide-linked.
CC -!- SIMILARITY: Belongs to the thrombospondin family.
CC -!- SIMILARITY: Contains 1 VWF domain.
CC -!- SIMILARITY: Contains 3 EGF-like domains.
CC -!- SIMILARITY: Contains 3 TSP type-1 domains.
CC -!- SIMILARITY: Contains 7 TSP type-3 domains.
CC -!- SIMILARITY: Contains 1 TSP N-terminal (TSPN) domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L04278; -; NOT_ANNOTATED_CDS.
CC HSP; P00740; IEDM.
CC InterPro; IPR001881; EGF Ca.
CC InterPro; IPR006209; EGF like.
CC InterPro; IPR006210; IEGF.
CC InterPro; IPR000884; TSP1.
CC InterPro; IPR008085; TSP1.
CC InterPro; IPR003367; tsp_1.
CC InterPro; IPR008859; TSP3.
CC InterPro; IPR003129; TSPN.
CC InterPro; IPR01007; VWF C.
CC Pfam; PF00008; EGF; 2.
CC Pfam; PF00090; tsp_1; 3.
CC Pfam; PF02412; tsp_3; 13.
CC Pfam; PF05735; TSPC; 1.
CC Pfam; PF02210; TSPN; 1.
CC Pfam; PF00093; VWC; 1.
CC PRINTS; PR01705; TSP1REPEAT.
CC SMART; SM00181; EGF; 2.
CC SMART; SM00203; TSP1; 3.
CC SMART; SM00210; TSPN; 1.
CC SMART; SM00214; VWC; 1.
CC PROSITE; PS00022; EGF 1; FALSE NEG.
CC PROSITE; PS01186; EGF 2; 1.
CC PROSITE; PS00026; EGF 3; 2.
CC PROSITE; PS50092; TSP1; 3.
CC PROSITE; PS01208; VWF; 1.
CC PROSITE; PS0184; VWF_2; 1.
CC Glycoprotein; Cell adhesion; Calcium-binding; Heparin-binding; Repeat;

DR InterPro: IPR000413; Integrin_alpha.
DR InterPro: IPR002035; VWF_A.
DR Pfam: PF01839; FG-GAP_3.
DR Pfam: PF00357; integrin_A; 1.
DR Pfam: PF00092; vwf; 1.
DR PRINTS: PR01185; INTEGRINA.
DR PRINTS: PR00453; VWFADOMAIN.
DR SMART: SM00191; Int alpha; 3.
DR SMART: SM00327; VWA; 1.
DR PROSITE: PS00242; INTEGRIN_ALPHA; 1.
DR PROSITE: PS00234; VWF; 1.
KW Integrin; Cell adhesion; Receptor; Glycoprotein; Transmembrane;
KW Signal; Repeat; Polymorphism; Magnesium; Calcium.
FT CHAIN 1 18
FT CHAIN 19 1179 INTEGRIN ALPHA-E.
FT CHAIN 19 1179 INTEGRIN ALPHA-E LIGHT CHAIN.
FT CHAIN 19 1177 INTEGRIN ALPHA-E HEAVY CHAIN.
FT CHAIN 173 1124 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 1125 1147 POTENTIAL.
FT TRANSMEM 1148 1179 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 1181 1198 GLU-RICH (ACIDIC).
FT DOMAIN 181 198
FT REPEAT 1 18
FT REPEAT 2 2
FT REPEAT 3 3
FT REPEAT 4 4
FT REPEAT 5 5
FT REPEAT 6 6
FT REPEAT 7 7
FT CA_BIND 522 530 POTENTIAL.
FT CA_BIND 586 594 POTENTIAL.
FT CA_BIND 654 662 POTENTIAL.
FT SITE 1150 1154 GPFKR MOTIF.
FT DISULFID 70 79 BY SIMILARITY.
FT DISULFID 126 159 BY SIMILARITY.
FT DISULFID 706 762 BY SIMILARITY.
FT DISULFID 823 829 BY SIMILARITY.
FT DISULFID 893 907 BY SIMILARITY.
FT DISULFID 1008 1033 BY SIMILARITY.
FT DISULFID 1041 1057 BY SIMILARITY.
FT CARBOHYD 49 49 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 726 726 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 782 782 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 857 857 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 934 934 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 954 954 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1065 1065 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1096 1096 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 360 360 D -> E.
FT VARIANT 1041 1041 /FTID=VAR_008884.
FT VARIANT 1041 1041 C -> S.
FT MUTAGEN 208 208 /FTID=VAR_008885.
FT MUTAGEN 316 316 D->A: LOSS OF E-CADHERIN BINDING.
FT CONFLICT 477 477 E->A: LOSS OF E-CADHERIN BINDING.
FT CONFLICT 482 482 V -> I (IN REF. 3).
FT CONFLICT 482 482 Q -> R (IN REF. 3).
FT CONFLICT 950 950 R -> W (IN REF. 3).
FT CONFLICT 1019 1019 A -> V (IN REF. 3).
SQ SEQUENCE 1179 AA; 13008 MW; E558902ED9D95E1 CRC64;
Query Match 0.7%; Score 8; DB 1; Length 1179;
Best Local Similarity 100.0%; Pred.No. 45;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 380 VLLGAVGA 387
DB 416 VLLGAVGA 423

RESULT 42
PMPC CHLMU STANDARD; PRT; 1460 AA.
AC 09DJY1; 2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable outer membrane protein pmpc precursor (Polymorphic membrane protein C).
DE PMPC OR TC0695.
GN Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MoPn / N199;
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidberg J.F., White O., Hickey E.K., Peterson J., Utterback T., Berry K., Bass S., Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R., Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L., Eisen J., Fraser C.M.;
RA "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39.";
RT Nucleic Acids Res. 28:1397-1406(2000).
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SUBCELLULAR LOCATION: Cell wall surface (elementary bodies) (Potential).
CC -!- SIMILARITY: Belongs to the PMP outer membrane protein family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC EMBL; AE002338; AAF39511.1; -.
DR PIR; D81675; D81675.
DR TIGS; TC0695; -.
DR InterPro: IPR006315; Autotransport.
DR InterPro: IPR003368; Chlamydia_PMP.
DR Pfam; PF02415; Chlamydia_PMP; 2.
DR TIGRFAMs; TIGR01414; autotrans_bar1; 1.
DR TIGRFAMs; TIGR01376; POMP repeat; 8.
KW Outer membrane; Signal; Multigene family; Complete proteome.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 1460 PROBABLE OUTER MEMBRANE PROTEIN PMPC.
SQ SEQUENCE 1460 AA; 154781 MW; C1F033E7907AC3AC CRC64;
Query Match 0.7%; Score 8; DB 1; Length 1460;
Best Local Similarity 100.0%; Pred.No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 583 SILKTPKQ 590
DB 1311 SILKTPKQ 1318
RESULT 43
SET2-CAEEL STANDARD; PRT; 1507 AA.
ID SET2-CAEEL
AC Q18221; Q950U6; Q950U7;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Protein set-2
GN SET-2 OR C36E6.9.
OC Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
CX NCBI_TaxID=6239;
RN [1]


```

DR InterPro; IPR007080; RNA_pol_Rpb1_1.
DR InterPro; IPR007066; RNA_pol_Rpb1_3.
DR InterPro; IPR007083; RNA_pol_Rpb1_4.
DR InterPro; IPR007081; RNA_pol_Rpb1_5.
DR InterPro; IPR006592; RNA_pol_N.
DR Pfam; PF04997; RNA_pol_Rpb1_1.
DR Pfam; PF00623; RNA_pol_Rpb1_2.
DR Pfam; PF04983; RNA_pol_Rpb1_3.
DR Pfam; PF05000; RNA_pol_Rpb1_4.
DR Pfam; PF04998; RNA_pol_Rpb1_5.
DR SMART; SM00563; RPOLA_N; 1.
DR Transferrase; DNA-directed RNA polymerase; Transcription; Zinc;
KW Zinc-finger; Nuclear protein.
FT ZN_FING 62 78 C2H2-TYPE (POTENTIAL).
FT CONFLICT 158 158 N->T (IN REF. 1).
SQ SEQUENCE 1664 AA; 186431 MW; DF65A7AA459D5E6D CRC64;

Query Match 0.7%; Score 8; DB 1; Length 1664;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 481 GSEITSVD 488
Db 8 GSEITSVD 15

RESULT 45
RRPL_BUNYW STANDARD; PRT; 2238 AA.
AC RRPL_BUNYW
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE RNA-directed RNA polymerase (EC 2.7.7.48) (L protein).
GN L.
OS Bunyamwera virus.
OC Viruses; ssRNA negative-strand viruses; Bunyaviridae; Orthobunyavirus.
OX NCBI_TaxID=35304;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90085791; PubMed=2596023;
RA Elliott R.M.;
RT "Nucleotide sequence analysis of the large (L) genomic RNA segment of Bunyamwera virus, the prototype of the family Bunyaviridae."
RL Virology 173:426-436(1989).
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA}[N].
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch)
CC
CC EMBL; X14383; CAA32553.1; -.
CC PIR; A33744; RRVUBY.
CC InterPro; IPR007322; Bunya RdRp.
CC InterPro; IPR007099; RNA_pol_NSvir.
CC Pfam; PF04196; Bunya RdRp; 1.
KW Transferrase; Nucleotidyltransferase; RNA-directed RNA polymerase.
SQ SEQUENCE 2238 AA; 258668 MW; IED00AB156BAC3DA CRC64;

Query Match 0.7%; Score 8; DB 1; Length 2238;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 184 FLINILKK 191
Db 1333 FLINILKK 1340

```

```

RESULT 46
LSTL_MOUSE
ID LSTL_MOUSE STANDARD; PRT; 95 AA.
AC O08843; O08844; Q9ZLH3, Q9ZL30;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Leukocyte specific transcript 1 protein (B144 protein).
OS LSTL OR B144.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RX TISSUE=B-cell;
RX MEDLINE=88031493; PubMed=3117682;
RA Tsuge I., Shen F.-W., Steinmetz M., Boyse E.A.;
RT "A gene in the H-2S.H-2D interval of the major histocompatibility complex which is transcribed in B cells and macrophages."
RL Immunogenetics 26:378-380(1987).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RX STRAIN=BALB/c; TISSUE=Macrophage;
RX MEDLINE=98035883; PubMed=9367684;
RA de Baey A., Fellerhoff B., Maier S., Martinuzzi S., Weidie U., Weiss E.H.;
RT "Complex expression pattern of the TNF region gene Lstl through differential regulation, initiation, and alternative splicing."
RL Genomics 45:591-600(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX Rowen L., Qin S., Madan A., Abbasi N., James R., Dickhoff R., Shafer T., Ratcliffe A., Loretz C., Lasky S., Hood L.;
RA "Sequence of the mouse major histocompatibility class III region."
RT Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION, SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RX MEDLINE=21372017; PubMed=11478849;
RA Raghunathan A., Sivakamasundari R., Wolenski J., Poddar R., Weissman S.M.;
RT "Functional analysis of B144/LSTl: a gene in the tumor necrosis factor cluster that induces formation of long filopodia in eukaryotic cells."
RL Exp. Cell Res. 268:230-244(2001).
CC -1- FUNCTION: Possible role in modulating immune responses. Has an inhibitory effect on lymphocyte proliferation. Induces morphological changes including production of filopodia and microspikes when overexpressed in a variety of cell types and may be involved in dendritic cell maturation.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Also detected in a perinuclear region corresponding to the localization of the Golgi apparatus and throughout the cytoplasm (By similarity).
CC -1- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=1; Synonyms=ml7r;
CC IsoId=O08843-1; Sequence=Displayed;
CC Name=2; Synonyms=ml1f;
CC IsoId=O08843-2; Sequence=VSP_050588;
CC -1- TISSUE SPECIFICITY: Expressed in spleen and at lower levels in thymus and liver.
CC -1- SIMILARITY: Belongs to the LSTl family.
CC -1- CAUTION: Ref.1 sequence differs from that shown due to a frameshift in position 58.
CC -1- CAUTION: Ref.3 sequence differs from that shown due to erroneous gene model prediction.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

```


OS *Shigella flexneri*.
 CC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 CC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=562, 83334, 623;
 RN [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 22-40.
 RP SPECIES=E.coli; STRAIN=K12;
 RC MEDLINE=93204894; PubMed=8455549;
 RX Yoshida T., Ueguchi C., Yamada H., Mizuno T.;
 RA "Function of the Escherichia coli nucleoid protein, H-NS: molecular
 RT analysis of a subset of proteins whose expression is enhanced in a
 RT hns deletion mutant.";
 RL Mol. Gen. Genet. 237:113-122(1993).
 RN [2]
 RN SEQUENCE FROM N.A.
 RP SPECIES=E.coli; STRAIN=K12 / MG1655;
 RC MEDLINE=94316500; PubMed=8041620;
 RX Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
 RA "Analysis of the Escherichia coli genome. V. DNA sequence of the
 RT region from 76.0 to 81.5 minutes.";
 RL Nucleic Acids Res. 22:2576-2586(1994).
 RN [3]
 RN SEQUENCE FROM N.A.
 RP SPECIES=E.coli; STRAIN=O157:H7 / EDL933 / ATCC 700927;
 RC MEDLINE=21074935; PubMed=11206551;
 RX Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
 RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
 RA Grotbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
 RA Welch R.A., Blattner F.R.;
 RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
 RL Nature 409:529-533(2001).
 RN [4]
 RN SEQUENCE FROM N.A.
 RP SPECIES=E.coli; STRAIN=O157:H7 / RIMD 0509952;
 RC MEDLINE=21156231; PubMed=11258796;
 RX Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe I.,
 RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
 RA Kuhara S., Shiba T., Hattori M., Shinagawa H.;
 RA "Complete genome sequence of enterohemorrhagic Escherichia coli
 RT O157:H7 and genomic comparison with a laboratory strain K-12.";
 RL DNA Res. 8:11-22(2001).
 RN [5]
 RN GENE NAME.
 RP SPECIES=E.coli;
 RC MEDLINE=94064579; PubMed=8244952;
 RX Yoshida T., Ueguchi C., Mizuno T.;
 RA "Physical map location of a set of Escherichia coli genes (hde) whose
 RT expression is affected by the nucleoid protein H-NS.";
 RL J. Bacteriol. 175:7747-7748(1993).
 RN [6]
 RN SEQUENCE OF 22-41.
 RP SPECIES=E.coli; STRAIN=K12 / W3110;
 RC MEDLINE=97443975; PubMed=9298646;
 RX Pasquali C., Sanchez J.-C., Ravier F., Golaz O., Hughes G.J.,
 RA Frutiger S., Paquet N., Wilkins M., Appel R.D., Bairoch A.,
 RA Hochstrasser D.F.;
 RL Submitted (SEP-1994) to Swiss-Prot.
 RN [7]
 RN SEQUENCE OF 22-33.
 RP SPECIES=E.coli; STRAIN=K12 / BMG2;
 RC MEDLINE=97443975; PubMed=9298646;
 RX Link A.J., Robison K., Church G.M.;
 RA "Comparing the predicted and observed properties of proteins encoded
 RT in the genome of Escherichia coli K-12.";
 RL Electrophoresis 18:1259-1313(1997).
 RN [8]
 RN SEQUENCE OF 22-31.
 RP SPECIES=E.coli; STRAIN=K12;
 RC MEDLINE=99085675; PubMed=9868784;
 RX Wasinger V.C., Humphery-Smith I.;
 RA "Small genes/gene-products in Escherichia coli K-12.";
 RT

FEMS Microbiol. Lett. 169:375-382(1998).
 [9]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.flexneri; STRAIN=301 / Serotype 2a;
 RX MEDLINE=22272406; PubMed=12384590;
 RA Jin Q., Yuan Z., Xu J., Wang Y., Shen Y., Lu W., Wang J., Liu H.,
 RA Yang J., Yang F., Zhang X., Zhang J., Yang G., Wu H., Qu D., Dong J.,
 RA Sun L., Xue Y., Zhao A., Gao Y., Zhu J., Kan B., Ding K., Chen S.,
 RA Cheng H., Yao Z., He B., Chen R., Ma D., Qiang B., Wen Y., Hou Y.,
 RA Yu J.;
 RA "Genome sequence of Shigella flexneri 2a: insights into pathogenicity
 RT through comparison with genomes of Escherichia coli K12 and O157.";
 RL Nucleic Acids Res. 30:4432-4441(2002).
 RN [10]
 RN SEQUENCE FROM N.A.
 RP SPECIES=S.flexneri; STRAIN=2457T / ATCC 700930 / Serotype 2a;
 RX MEDLINE=22590274; PubMed=12704152;
 RA Wei J., Goldberg M.B., Burland V., Venkatesan M.M., Deng W.,
 RA Pournier G., Mayhew G.F., Plunkett G. III, Rose D.J., Darling A.,
 RA Mau B., Perna N.T., Payne S.M., Runyen-Janecky L.J., Zhou S.,
 RA Schwartz D.C., Blattner F.R.;
 RA "Complete genome sequence and comparative genomics of Shigella
 RT flexneri serotype 2a strain 2457T.";
 RL Infect. Immun. 71:2775-2786(2003).
 RN [11]
 RN MASS SPECTROMETRY.
 RP SPECIES=E.coli; STRAIN=ATCC 15597;
 RC MEDLINE=22067786; PubMed=12071744;
 RX Reid G.E., Shang H., Hogan J.M., Lee G.U., McLuckey S.A.;
 RA "Gas-phase concentration, purification, and identification of whole
 RT proteins from complex mixtures.";
 RL J. Am. Chem. Soc. 124:7353-7362(2002).
 RN [12]
 RN X-RAY CRYSTALLOGRAPHY.
 RP SPECIES=E.coli;
 RC MEDLINE=98400493; PubMed=9731767;
 RX Yang F., Gustafson K.R., Boyd M.R., Wlodawer A.;
 RA "Crystal structure of Escherichia coli HdeA.";
 RL Nat. Struct. Biol. 5:763-764(1998).
 RN [13]
 RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS).
 RP SPECIES=E.coli;
 RC MEDLINE=20090957; PubMed=10623550;
 RX Gajiwala K.S., Burley S.K.;
 RA "HDEA, a periplasmic protein that supports acid resistance in
 RT pathogenic enteric bacteria.";
 RL J. Mol. Biol. 295:605-612(2000).
 CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).
 CC -1- MASS SPECTROMETRY: MW=9742; METHOD=Electrospray; RANGE=22-110.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to licenses@isb-sib.ch).
 CC -----
 DR ENBL; D11109; BAA01883.1; -;
 DR ENBL; U00039; AAB18486.1; -;
 DR ENBL; AE000427; AAC76535.1; -;
 DR ENBL; AE005576; AAG58651.1; -;
 DR ENBL; AF002565; BAB37813.1; -;
 DR ENBL; AE015362; AAN44999.1; -;
 DR ENBL; AE016992; AAP19187.1; -;
 DR PIR; F91177; F91177;
 DR PIR; G86023; G86023;
 DR PIR; S30268; S30268;
 DR PDB; 1DJS; 10-DEC-99.
 DR PDB; 1BGS; 16-SEP-98.
 DR SWISS-2DPAGE; P26604; COLI.
 DR EcoGene; EG11398; hdeA.
 DR Periplasmic; Signal; 3D-structure; Complete proteome.
 KW

```

FT SIGNAL 1 21
FT CHAIN 22 110
FT DISULFID 39 87
FT HELIX 34 36
FT STRAND 38 38
FT HELIX 39 43
FT TURN 44 44
FT HELIX 47 49
FT HELIX 50 60
FT TURN 61 64
FT HELIX 66 68
FT HELIX 73 88
FT TURN 89 89
FT TURN 91 92
FT STRAND 94 94
FT HELIX 95 107
SQ SEQUENCE 110 AA; 11858 MW; 063262C4863FA2E9 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGILLLL 1156
DB 9 LGGILLLL 15

RESULT 49
INS1_RAT STANDARD; PRT; 110 AA.
AC P01322;
AT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin 1 precursor.
GN INS1 OR INS-1
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RX SEQUENCE FROM N.A.
RP MEDLINE=80045034; PubMed=498283;
RA Cordell B., Bell G.I., Fischer E., Denoto F.M., Ullrich A.,
RA Pictet R.L., Rutter W.J., Goodman H.M.;
RT "Isolation and characterization of a cloned rat insulin gene.";
RL Cell 18:533-543 (1979).
RN [2]
RX SEQUENCE FROM N.A.
RP STRAIN=Sprague-Dawley; Tissue=Liver;
RX MEDLINE=80045035; PubMed=498284;
RA Lomedico P., Rosenthal N., Efstratiadis A., Gilbert W., Kolodner R.,
RA Tizard R.;
RT "The structure and evolution of the two nonallelic rat preproinsulin
genes.";
RL Cell 18:545-558 (1979).
RN [3]
RX SEQUENCE FROM N.A.
RP MEDLINE=80240379; PubMed=6249167;
RA Lomedico P.T., Rosenthal N., Kolodner R., Efstratiadis A.,
RA Gilbert W.;
RT "The structure of rat preproinsulin genes.";
RL Ann. N.Y. Acad. Sci. 343:425-432 (1980).
RN [4]
RX SEQUENCE OF 25-54 AND 90-110.
RP MEDLINE=70067613; PubMed=4311938;
RA Steiner D.F., Clark J.L., Nolan C., Rubenstein A.H., Margoliash E.,
RA Aten B., Oyer P.E.;
RT "Proinsulin and the biosynthesis of insulin.";
RL Recent Prog. Horm. Res. 25:207-282 (1969).
RN [5]
RX SEQUENCE OF 57-87.
RP STRAIN=New Zealand white; Tissue=Pancreas;
RP MEDLINE=73061498; PubMed=4640931;

```

```

RA Tager H.S., Steiner D.F.;
RT "Primary structures of the proinsulin connecting peptides of the rat
and the horse.";
RL J. Biol. Chem. 247:7936-7940 (1972).
RN [6]
RP SEQUENCE OF 57-87, AND REVISIONS.
RX MEDLINE=72177385; PubMed=4554104;
RA Markussen J., Sundby F.;
RT "Rat-proinsulin C-peptides. Amino-acid sequences.";
RL Eur. J. Biochem. 25:153-162 (1972).
CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
increases cell permeability to monosaccharides, amino acids and
fatty acids. It accelerates glycolysis, the pentose phosphate
cycle, and glycogen synthesis in liver.
CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
disulfide bonds.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; V01242; CAA24559.1; -
DR EMBL; J00747; AAA41442.1; -
DR EMBL; M25584; AAA41439.1; -
DR PIR; A90788; IPRT1.
DR HSP; P01308; 1A7F.
DR InterPro; IP0004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; ILGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
DR INSULIN family; Hormone; Glucose metabolism; Signal; Multigene family.
FT SIGNAL 1 24
FT CHAIN 25 54
FT PROPEP 57 87
FT CHAIN 90 110
FT DISULFID 31 96
FT DISULFID 43 109
FT DISULFID 95 100
SQ SEQUENCE 110 AA; 12420 MW; 51D606DA54AE3533 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVL 1161
DB 10 LLALLVL 16

RESULT 50
INS_RABIT STANDARD; PRT; 110 AA.
AC P01311;
AT 21-JUL-1986 (Rel. 01, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin precursor.
GN INS.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RX SEQUENCE FROM N.A.
RP STRAIN=New Zealand white; Tissue=Pancreas;
RP MEDLINE=94179230; PubMed=8132571;

```



```

RA Devaskar S.U., Giddings S.J., Rajakumar P.A., Carnaghi L.R.,
RA Menon R.K., Zahm D.S.;
RT "Insulin gene expression and insulin synthesis in mammalian neuronal
RT cells.";
RL J. Biol. Chem. 269:8445-8454(1994).
RN [2]
RP SEQUENCE OF 25-54 AND 90-110.
RX MEDLINE=66160119; PubMed=5949593;
RA Smith L.F.;
RT "Species variation in the amino acid sequence of insulin.";
RL Am. J. Med. 40:662-666(1966).
RN [3]
RP SEQUENCE OF 56-110 FROM N.A.
RA Giddings S.J., Carnaghi L.R., Devaskar S.U.;
RL Submitted (APR-1991) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Insulin decreases blood glucose concentration. It
CC increases cell permeability to monosaccharides, amino acids and
CC fatty acids. It accelerates glycolysis, the pentose phosphate
CC cycle, and glycogen synthesis in liver.
CC -!- SUBUNIT: Heterodimer of a B chain and an A chain linked by two
CC disulfide bonds.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; U03610; AAA19033.1; -.
DR EMBL; M61153; AAA17540.1; -.
DR PIR; A53438; INRB.
DR HSP; P01308; ILYM.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; ILGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Hormone; Glucose metabolism; Signal.
FT SIGNAL 1 24
FT CHAIN 25 54 INSULIN B CHAIN.
FT PROPEP 57 87 C PEPTIDE.
FT CHAIN 90 110 INSULIN A CHAIN.
FT DISULFID 31 96 INTERCHAIN.
FT DISULFID 43 109 INTERCHAIN.
FT DISULFID 95 100
FT CONFLICT 83 83 E -> Y (IN REF. 3).
SQ SEQUENCE 110 AA; 11938 MW; 82D2975B85D77FA8 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1155 LLALLVL 1161
Db 10 LLALLVL 16

RESULT 51
IGF_MXGL
ID -IGF_MXGL STANDARD; PRT; 139 AA.
AC P22618;
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Insulin-like growth factor precursor (IGF) (Fragment).
OS Myxine glutinosa (Atlantic hagfish).
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;
OC Myxiniidae; Myxinae; Myxine.
NCBI_TaxID=7769;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91115860; PubMed=1989990;
RA Nagamatsu S., Chan S.J., Falkner S., Steiner D.F.;
RT "Evolution of the insulin gene superfamily. Sequence of a
RT preproinsulin-like growth factor cDNA from the Atlantic hagfish.";
RL J. Biol. Chem. 266:2397-2402(1991).
CC -!- FUNCTION: The insulin-like growth factors, isolated from plasma,
CC are structurally and functionally related to insulin but have a
CC much higher growth-promoting activity.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: Belongs to the insulin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; M57735; AAA49265.1; -.
DR PIR; A39612; A38612.
DR HSP; P01344; IGF2.
DR InterPro; IPR004825; Ins/IGF/relax.
DR Pfam; PF00049; Insulin; 1.
DR PRINTS; PR00277; INSULINB.
DR SMART; SM00078; ILGF; 1.
DR PROSITE; PS00262; INSULIN; 1.
KW Insulin family; Mitogen; Growth factor; Signal.
FT NON_TER 1 1
FT SIGNAL <1 38
FT CHAIN 39 139 INSULIN-LIKE GROWTH FACTOR.
FT DOMAIN 39 67 B.
FT DOMAIN 68 82 C.
FT DOMAIN 83 103 A.
FT DOMAIN 104 113 D.
FT DOMAIN 114 139 E.
SQ SEQUENCE 139 AA; 16087 MW; 2FC888C8D074FAC1 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 139;
Best Local Similarity 100.0%; Pred. No. 77;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALL 1159
Db 28 LLLALL 34

RESULT 52
LAMP_PETWA
ID LAMP_PETWA STANDARD; PRT; 139 AA.
AC P33575; P33576;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE Lamprin 0.9 precursor (Cartilage matrix protein).
OS Petromyzon marinus (Sea lamprey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Hyperoartia;
OC Petromyzontiformes; Petromyzontidae; Petromyzon.
NCBI_TaxID=7757;
RN [1]_TaxID=7757;
RP SEQUENCE FROM N.A. (ISOFORMS 0.9-10 AND 0.9-12), AND SEQUENCE OF
RP 20-44.
RC TISSUE=Cartilage;
RX MEDLINE=93123269; PubMed=7678258;
RA Robson P., Wright G.M., Sitarz E., Maiti A., Rawat M., Youson J.H.,
RA Keeley F.W.;
RT "Characterization of lamprin, an unusual matrix protein from lamprey
RT cartilage. Implications for evolution, structure, and assembly of
RT elastin and other fibrillar proteins.";
RL J. Biol. Chem. 268:1440-1447(1993).
CC -!- FUNCTION: Self-aggregating protein that makes part of the soluble

```

```

CC form of lamprin.
CC -!- SUBUNIT: The polymeric lamprin chains self-aggregate to form
CC fibers and have secondary structures particularly rich in beta-
CC sheets and in beta-turns.
CC -!- SUBCELLULAR LOCATION: Extracellular matrix.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=2;
CC Name=0.9-12;
CC IsoId=P33575-1; Sequence=Displayed;
CC Name=0.9-10;
CC IsoId=P33575-2; Sequence=VSP_004302;
CC Note=No experimental confirmation available;
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; L05925; AAA49269.1; --
CC EMBL; L05924; AAA49268.1; --
CC PIR; C45051; C45051.
CC Repeat; Connective tissue; Signal; Alternative splicing.
CC SIGNAL 1 19 LAMPRIN 0.9.
CC CHAIN 20 139
CC DOMAIN 42 110 8 X 5 AA APPROXIMATE REPEATS.
CC REPEAT 42 46 1.
CC REPEAT 47 51 2.
CC REPEAT 52 56 3.
CC REPEAT 57 61 4.
CC REPEAT 62 66 5.
CC REPEAT 67 71 6.
CC REPEAT 92 96 7.
CC REPEAT 106 110 8.
CC VARSPLIC 86 104 Missing (in isoform 0.9-10).
CC /FTId=VSP_004302.
CC SEQUENCE 139 AA; 13257 MW; E248ABB7A6CEC7C3 CRC64;
CC
CC Query Match 0.6%; Score 7; DB 1; Length 139;
CC Best Local Similarity 100.0%; Pred. No. 77;
CC Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC
CC QY 1157 ALLVIAL 1163
CC Db 7 ALLVIAL 13
CC
CC RESULT 53
CC HEAD CHICK
CC ID HEAD_CHICK STANDARD; PRT; 141 AA.
CC AC P02001;
CC DT 21-JUL-1986 (Rel. 01, Created)
CC DT 21-JUL-1986 (Rel. 01, Last sequence update)
CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
CC DE Hemoglobin alpha-D chain.
CC GN HBA
CC OS Gallus gallus (Chicken).
CC CC Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC CC Archosauria; Aves; Neognathae; Galliformes; Phasianinae;
CC CC Gallus.
CC OX NCBI_TaxID=9031;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC STRAIN=White leghorn;
CC RX MEDLINE=92020223; PubMed=1656392;
CC RA Lewis W., Lee J.D., Dodgson J.B.;
CC RT "Adult chicken alpha-globin gene expression in transfected Q76 quail
CC cells: evidence for a negative regulatory element in the alpha D gene
CC region.";
CC RL Nucleic Acids Res. 19:5321-5329(1991).
CC RN [2]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=83161047; PubMed=6300093;
RA Dodgson J.B., Engel J.D.;
RT "The nucleotide sequence of the adult chicken alpha-globin genes.";
RL J. Biol. Chem. 258:4623-4629(1983).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=82082384; PubMed=6273837;
RA Dodgson J.B., McCune K.C., Rusling D.J., Krust A., Engel J.D.;
RT "Adult chicken alpha-globin genes alpha A and alpha D: no anemic
RT shock alpha-globin exists in domestic chickens.";
RL Proc. Natl. Acad. Sci. U.S.A. 78:5998-6002(1981).
RN [4]
RP PRELIMINARY SEQUENCE.
RX MEDLINE=76189993; PubMed=1225908;
RA Takei H., Ota Y., Wu K.C., Kiyohara T., Matsuda G.;
RT "Amino acid sequence of the alpha chain of chicken AI hemoglobin.";
RL J. Biochem. 77:1345-1347(1975).
RN [5]
RP SEQUENCE OF 1-63 AND 94-122.
RC TISSUE=Embryo;
RX MEDLINE=82098109; PubMed=7054172;
RA Chapman B.S., Hood L.E., Tobin A.J.;
RT "Minor early embryonic chick hemoglobin M. Amino acid sequences of
RT the epsilon and alpha D chains.";
RL J. Biol. Chem. 257:651-658(1982).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX MEDLINE=95156923; PubMed=10037733;
RA Knapp J.E., Oliveira M.A., Xie Q., Ernst S.R., Riggs A.F.,
RA Hackert M.S.;
RT "The structural and functional analysis of the hemoglobin D component
RT from chicken.";
RL J. Biol. Chem. 274:6411-6420(1999).
CC -!- FUNCTION: Involved in oxygen transport from the lung to the
CC various peripheral tissues.
CC -!- SUBUNIT: Heterotetramer of two alpha-D chains and two beta chains.
CC The component D forms dimers of tetramers upon deoxygenation.
CC -!- TISSUE SPECIFICITY: Red blood cells.
CC -!- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor
CC hemoglobin component, called hemoglobin d, which is expressed in
CC late embryonic and adult life.
CC -!- MISCELLANEOUS: REF.5 CHAIN WAS ISOLATED FROM HEM, THE LEAST
CC ABUNDANT OF THE FOUR EARLY CHICK HEMOGLOBINS.
CC -!- SIMILARITY: Belongs to the globin family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X59989; CAA42605.1; --
CC EMBL; V00411; CAA23702.1; --
CC EMBL; M15378; AAA48584.1; ALT_SEQ.
CC EMBL; J00853; AAA48800.1; --
CC PIR; B92421; HACH1.
CC PDB; 1HBR; 26-MAR-99.
CC InterPro; IPR002336; Alpha_haem.
CC InterPro; IPR000971; Globin.
CC Pfam; PF00042; globin; 1.
CC PRINTS; PR00612; ALPHAHAE.
CC PROSITE; PS01033; GLOBIN; 1.
CC Heme; Oxygen transport; transport; Erythrocyte;
CC 3D-structure.
CC FT METAL 58 58 IRON (HEME DISTAL LIGAND).
CC FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
CC FT CONFLICT 16 16 K -> R (IN REF. 1).
CC FT CONFLICT 107 107 V -> C (IN REF. 3 AND 5).
CC FT HELIX 4 17
CC FT TURN 18 19

```

```
PT HELIX 21 35
PT HELIX 37 42
FT TURN 44 51
FT TURN 50 51
FT HELIX 53 71
FT TURN 72 74
FT HELIX 76 79
FT HELIX 81 88
FT TURN 89 91
FT TURN 95 95
FT HELIX 96 113
FT HELIX 114 116
FT HELIX 119 136
FT TURN 137 137
SQ SEQUENCE 141 AA; 15695 MW; 1FE42696B7B5384 CRC64;

Query Match 0.8%; Score 7; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 625 GAGNAV 631
Db 64 GAGNAV 70

RESULT 54
HEAD MELGA STANDARD; PRT; 141 AA.
AC P81024;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hemoglobin alpha-D chain.
GN HEAD.
OS Meleagris gallopavo (Common turkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Meleagris.
CX NCBI_TaxID=9103;
RN [1]
RP SEQUENCE.
RX MEDLINE=96066279; PubMed=7576241;
RA Eguchi Y., Ikehara T., Kayo S., Eguchi T., Takei H.;
RT "Amino acid sequence of alpha- and beta-polypeptide chains of turkey
(RL Meleagris gallopavo) hemoglobin.";
RL Biol. Chem. Hoppe-Seyler 376:437-440 (1995).
CC -!- FUNCTION: Involved in oxygen transport from the lung to the
various peripheral tissues.
CC -!- SUBUNIT: Heterotetramer of two alpha-D chains and two beta chains.
CC -!- TISSUE SPECIFICITY: Red blood cells.
CC -!- DEVELOPMENTAL STAGE: In birds, the alpha-D chain occurs in a minor
late embryonic component, called hemoglobin d, which is expressed in
CC -!- SIMILARITY: Belongs to the globin family.
DR HSP; P02001; IHRB.
DR HSP; P02001; IHRB.
DR InterPro; IPR002338; Alpha haem.
DR InterPro; IPR000971; Globin.
DR Pfam; PF00042; globin; 1.
DR PRINTS; PS00612; ALPHAHAE.
DR PROSITE; PS01033; GLOBIN; 1.
KW Heme; Oxygen transport; Transport; Erythrocyte.
FT METAL 58 58 IRON (HEME DISTAL LIGAND).
FT METAL 87 87 IRON (HEME PROXIMAL LIGAND).
SQ SEQUENCE 141 AA; 15665 MW; 59942696887954E7 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 141;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 625 GAGNAV 631
Db 64 GAGNAV 70
```

```
RESULT 55
MK_CHICK
ID MK_CHICK STANDARD; PRT; 142 AA.
AC P24052;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Midkine precursor (Retinoic acid-induced heparin-binding protein)
DE (RI-HB).
GN RIHB.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
OC Gallus.
CX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 80-85; 112-118 AND 121-133.
RX TISSUE=Embryo;
RX MEDLINE=91207359; PubMed=2018506;
RA Urios P., Duprez D., le Caer J.-P., Courtois Y., Vigny M., Laurent M.;
RT "Molecular cloning of RI-HB, a heparin binding protein regulated by
retinoic acid.";
RL Biochem. Biophys. Res. Commun. 175:617-624 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95010085; PubMed=7925417;
RA Duprez D., Treaguer J., Pecqueur C., Vigny M.R.;
RT "Organisation and promoter activity of the
retinoic-acid-induced-heparin-binding (RIHB) gene.";
RN [3]
RP SEQUENCE OF 22-77.
RX MEDLINE=91128406; PubMed=1993066;
RA Raulais D., Lagente-Chevallier O., Guettet C., Duprez D., Courtois Y.,
Vigny M.;
RT "A new heparin binding protein regulated by retinoic acid from chick
embryo.";
RL Biochem. Biophys. Res. Commun. 174:708-715 (1991).
RN [4]
RP SEQUENCE OF 22-77.
RX STRAIN=White leghorn; TISSUE=Embryo;
RX MEDLINE=90108010; PubMed=2558016;
RA Vigny M., Raulais D., Puzenat N., Duprez D., Hartman M.P.,
Jeanny J.C., Courtois Y.;
RT "Identification of a new heparin-binding protein localized within
chick basement membranes.";
RL Eur. J. Biochem. 186:733-740 (1989).
CC -!- FUNCTION: Has mitogenic activity, and neurite extension activity
for PC12 cells.
CC -!- SUBCELLULAR LOCATION: BASEMENT MEMBRANES IN EARLY EMBRYONIC
TISSUES, AND CELL SURFACE OF NEUROECTODERMAL CELLS.
CC -!- DEVELOPMENTAL STAGE: Essentially expressed during embryogenesis.
CC -!- INDUCTION: By retinoic acid.
CC -!- SIMILARITY: Belongs to the pleiotrophin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC EMBL; M61754; -; NOT ANNOTATED_CDS.
DR EMBL; X76482; CAA54020.1; -.
DR PIR; JT0573; JT0573.
DR HSP; P21741; IMKC.
DR InterPro; IPR000762; PTN MK.
DR Pfam; PF01091; PTN MK C; 1.
DR Pfam; PF05196; PTN MK N; 1.
DR PRINTS; PR00269; PTNWDKINE.
DR ProDom; PD005592; PTN MK; 1.
```

DR SMART: SMO0193; PTN: 1.
 DR PROSITE: PS00619; PTN_MK_1; 1.
 DR PROSITE: PS00620; PTN_MK_2; 1.
 KW Growth factor; Mitogen; Differentiation; Heparin-binding; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 142 MIDKINE.
 FT DISULFID 36 60 BY SIMILARITY.
 FT DISULFID 44 69 BY SIMILARITY.
 FT DISULFID 51 73 BY SIMILARITY.
 FT DISULFID 83 115 BY SIMILARITY.
 FT DISULFID 93 125 BY SIMILARITY.
 FT CONFLICT 89 93 S -> R (IN REF. 2).
 FT CONFLICT 93 93 C -> G (IN REF. 2).
 SQ SEQUENCE 142 AA; 1557 MW; 9D05CAF9B558451B CRC64;
 Query Match 0.6%; Score 7; DB 1; Length 142;
 Best Local Similarity 100.0%; Pred. No. 78;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1153 LLLALL 1159
 Db 7 LLLALL 13
 RESULT 56
 BIK_HUMAN
 ID BIK_HUMAN STANDARD; PRT; 160 AA.
 AC Q1323; Q16582;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Bcl-2 interacting killer (Apoptosis inducer NBK) (BP4) (BIP1).
 GN BIK OR NBK.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 [1]
 RN SEQUENCE FROM N.A.
 RC TISSUE=B-cell;
 RX MEDLINE=96068922; PubMed=7478623;
 RA Boyd J.M., Gallo G.J., Elangovan B., Houghton A.B., Malstrom S.,
 RA Avery B.J., Ebb R.G., Subramanian T., Chittenden T., Lutz R.J.,
 RA Chinnadurai G.
 RT "Bik" a novel death-inducing protein shares a distinct sequence motif
 RT with Bcl-2 family proteins and interacts with viral and cellular
 RT survival-promoting proteins."
 RL Oncogene 11:1921-1928(1995).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=96413338; PubMed=8816500;
 RA Han J., Sabbatini P., White E.;
 RT "Induction of apoptosis by human Nbk/Bik, a BH3-containing protein
 RT that interacts with E1B 19K."
 RL Mol. Cell. Biol. 16:5857-5864(1996).
 RN [3]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Lymphoid;
 RA Pun K.-T., Farrow S.N., Raven T., Wride C.J., White J.H.M., Brown R.;
 RL Submitted (JUL-1995) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=99431905; PubMed=10500065;
 RA Castell A., Ino Y., Louis D.N., Ramesh V., Gussella J.F., Rustgi A.K.;
 RT "Mapping of a target region of allelic loss to a 0.5-cM interval on
 RT chromosome 22q13 in human colorectal cancer."
 RL Gastroenterology 117:831-837(1999).
 RN [5]
 RN SEQUENCE FROM N.A.
 RA Rieder M.J., Livingston R.J., Daniels M.R., Chung M.-W.,
 RA Miyamoto K.E., Nguyen C.P., Nguyen D.A., Poel C.L., Robertson P.D.,
 RA Schackwitz W.S., Sherwood J.K., Witrak L.A., Nickerson D.A.;
 RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.

RP SEQUENCE FROM N.A.
 RX MEDLINE=20057165; PubMed=10591208;
 RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,
 RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
 RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
 RA Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
 RA Burill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
 RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
 RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
 RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
 RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
 RA Gilbert J.G.R., Goward M.E., Graffham D.V., Griffiths M.N.D., Hall C.,
 RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
 RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
 RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
 RA Martyn I.D., Mashrafi-Mohammadi M., Matthews L.H., McCann O.T.,
 RA McElay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
 RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.,
 RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
 RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
 RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
 RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
 RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
 RA Wright C.B., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
 RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
 RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyama S.,
 RA Roe B.A., Chen P., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
 RA Lewis J., Lewis S., Lin S.-P., Loh P., Hua A., Kenton S., Lai H., Lao H.I.,
 RA Devis J., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
 RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
 RA Zhang M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
 RA Fulton R., Johnson D., Benis G., Bentley D., Bradshaw H., Bourne S.,
 RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
 RA Hinds K., Kemp K., Latreille P., Layman D., Ozerky P., Rohlfing T.,
 RA Scheet P., Walker C., Wamsley A., Wohldmann P., Pepin K., Nelson J.,
 RA Koef I., Bedell J.A., Hillier L.W., Mardis E., Waterston R.,
 RA Wilson R., Emanuel B.S., Shaikh T., Kurahashi H., Saitta S.,
 RA Sudar M.L., McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E.,
 RA Edlmann L., Kim U.J., Shizuya H., Simon M.I., Dumanski J.P.,
 RA Peyrard M., Keda D., Seroussi E., Franssen I., Tapia I., Bruder C.E.,
 RA O'Brien K.P., Wilkinson P., Bodenteich A., Hartman K., Hu X.,
 RA Khan A.S., Lane L., Tilahun Y., Wright H.;
 RT "The DNA sequence of human chromosome 22."
 RL Nature 402:489-495(1999).
 RN [7]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Lymph;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raba S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Madan A., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Fahy J., Heiton E., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RN MUTAGENESIS, AND FUNCTION OF BH3 DOMAIN.
 RP MEDLINE=9601131; PubMed=8521816;
 RX

Chittenden T., Flemington C., Houghton A.B., Ebb R.G., Gallo G.J.,
Elangovan B., Chinnadurai G., Lutz R.J.;
"A conserved domain in Bak, distinct from BH1 and BH2, mediates cell
death and protein binding functions.";
EMBO J. 14:5589-5596(1995)
CC -!- FUNCTION: Accelerates programmed cell death. Binding to the
CC apoptosis repressors Bcl-X(L), BHRFl, Bcl-2 or its adenovirus
CC homolog E1B 19k protein suppresses this death-promoting activity.
CC Does not interact with BAX.
CC -!- SUBCELLULAR LOCATION: Around the nuclear envelope, and in
CC cytoplasmic membranes.
CC -!- DOMAIN: Intact BH3 domain is required by BIK, BID, BAK, BAD and
CC with anti-apoptotic members of the Bcl-2 family.
CC -!- SIMILARITY: Contains 1 Bcl-2 homology 3 (BH3) domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U34584; AAC50413.1; -;
CC EMBL; U49730; AAC79124.1; -;
CC EMBL; X89986; CAA62013.1; -;
CC EMBL; AF174424; AAF01156.1; -;
CC EMBL; AF174421; AAF01156.1; JOINED.
CC EMBL; AF174422; AAF01156.1; JOINED.
CC EMBL; AF174423; AAF01156.1; JOINED.
CC EMBL; AV245248; AAF061089.1; -;
CC EMBL; AL022237; CAA18260.2; -;
CC EMBL; BC001599; AAO01599.1; -;
CC EMBL; PIR; S58214; S58214.
CC EMBL; Genew; HGNC:1051; BIK.
CC MIM; 603392; -;
CC GO; GO:0008632; P:apoptotic program; TAS.
CC GO; GO:0006917; P:induction of apoptosis; TAS.
CC InterPro; IPR000712; Bcl2_BH.
CC PROSITE; PS01259; BH3; 1.
CC Apoptosis; Transmembrane.
CC DOMAIN 57 71
CC TRANSMEM 136 156
CC DOMAIN 137 158
CC CONFLICT 149 150 PL -> LP (IN REF. 1).
CC SEQUENCE 160 AA; 18016 MW; 89034F4443F5A136 CRC64;
Query Match 0.6%; Score 7; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1153 LLLIALL 1159
DB 141 LLLALL 147
RESULT 57
NUOE AQUAE
ID NUOE AQUAE STANDARD; PRT; 160 AA.
AC O66842;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE NADH-quinone oxidoreductase chain E (EC 1.6.99.5) (NADH dehydrogenase
DE I, chain E) (NDH-1, chain E).
GN NUOE OR AQ 574.
OS Aquifex aeolicus.
OC Bacteria; Aquificae; Aquificales; Aquificaceae; Aquifex.
OX NCBI_TaxID=63363;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VF5;

MEDLINE=98196666; PubMed=9537320;
Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
Graham D.E., Overbeek R., Shear M.A., Keller M., Aujay M., Huber R.,
Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
"The complete genome of the hyperthermophilic bacterium Aquifex
aeolicus.";
Nature 392:353-358(1998).
CC -!- FUNCTION: NDH-1 shuttles electrons from NADH, via FMN and iron-
CC sulfur (Fe-S) centers, to quinones in the respiratory chain.
CC Couples the redox reaction to proton translocation (for every two
CC electrons transferred, four hydrogen ions are translocated across
CC the cytoplasmic membrane), and thus conserves the redox energy in
CC a proton gradient (By similarity).
CC -!- CATALYTIC ACTIVITY: NADH + quinone = NAD(+) + quinol.
CC -!- COFACTOR: Binds 1 2Fe-2S cluster (Potential).
CC -!- SIMILARITY: Belongs to the complex I 24 kDa subunit family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000696; AAC06799.1; -;
CC PIR; F70351; F70351.
CC InterPro; IPR002023; Cmplx1_24kDa.
CC Pfam; PF01257; complex_24kD; 1.
CC PRODOM; PD003859; Cmplx1_24kDa; 1.
CC PROSITE; PS01099; COMPLEX_24k; 1.
CC Oxidoreductase; NAD; Quinone; Metal-binding; Iron-sulfur; Iron;
CC 2Fe-2S; Complete proteome.
CC METAL 86 86
CC METAL 91 91
CC METAL 127 127
CC METAL 131 131
CC SEQUENCE 160 AA; 18550 MW; 425D81995A491B35 CRC64;
Query Match 0.6%; Score 7; DB 1; Length 160;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 413 BPEELK 419
DB 7 BPEELK 13
RESULT 58
TCH2 ARATH
ID TCH2 ARATH STANDARD; PRT; 161 AA.
AC P25070; O22592;
DT 01-MAY-1992 (Rel. 22, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Calmodulin-related protein 2, touch-induced.
GN TCH2 OR ATSG3770 OR K22F20.1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Johnson K.A., Braam J.;
RA Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=98087489; PubMed=9872454;
RA Nakamura Y., Sato S., Asamizu E., Kaneko T., Kotani H., Miyajima N.,
RA Tabata S.;

RT "Structural analysis of Arabidopsis thaliana chromosome 5. VII.
 RT Sequence features of the regions of 1,013,767 bp covered by sixteen
 RT physically assigned p1 and TAC clones."
 RL DNA Res. 5:297-308(1998).
 RN [3]
 RP SEQUENCE OF 26-70 FROM N.A.
 RC STRAIN=cv. Columbia;
 RA MEDLINE=90150263; PubMed=2302732;
 RX Braam J., Davis R.W.;
 RT "Rain-, wind-, and touch-induced expression of calmodulin and
 RT calmodulin-related genes in Arabidopsis."
 RL Cell 60:357-364(1990).
 RN [4]
 RP 3D-STRUCTURE MODELING OF 7-158.
 RC MEDLINE=97189489; PubMed=9037719;
 RA Khan A.R., Johnson K.A., Braam J., James M.N.G.;
 RX "Comparative modeling of the three-dimensional structure of the
 RT calmodulin-related TCH2 protein from Arabidopsis."
 RL Proteins 27:144-153(1997).
 CC -!- INDUCTION: By rain-, wind-, and touch (thigmomorphogenesis).
 CC -!- SIMILARITY: Contains 4 EF-hand calcium-binding domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; AF026473; BAB2713.1; -
 CC PIR; AB016873; BAB10353.1; -
 CC PDB; C34569; C34669.
 CC
 CC InterPro; IPR003293; Calflagin.
 CC InterPro; IPR002048; EF-hand.
 CC Pfam; PF00036; ehand; 4.
 CC PRINTS; PR01362; CALFLAGIN.
 CC ProDom; PD000012; EF-hand; 2.
 CC SMART; SM00054; EFH; 4.
 CC PROSITE; PS00018; EF_HAND; 4.
 CC Calcium-binding; Repeat; 3D-structure.
 CC CA_BIND 26 37 EF-HAND 1 (POTENTIAL).
 CC CA_BIND 62 73 EF-HAND 2 (POTENTIAL).
 CC CA_BIND 103 114 EF-HAND 3 (POTENTIAL).
 CC CA_BIND 139 150 EF-HAND 4 (POTENTIAL).
 CC CONFLICT 54 54
 CC HELIX 12 25
 CC TURN 27 28
 CC STRAND 33 33
 CC HELIX 35 44
 CC TURN 45 45
 CC HELIX 51 61
 CC TURN 63 64
 CC STRAND 69 69
 CC HELIX 71 102
 CC TURN 104 105
 CC STRAND 110 110
 CC HELIX 112 121
 CC TURN 122 123
 CC HELIX 128 138
 CC STRAND 146 146
 CC HELIX 148 156
 CC SEQUENCE 161 AA; 17546 MW; A7324A5C0B0CBDB9 CRC64;
 KW
 FT CA_BIND 26 37 EF-HAND 1 (POTENTIAL).
 FT CA_BIND 62 73 EF-HAND 2 (POTENTIAL).
 FT CA_BIND 103 114 EF-HAND 3 (POTENTIAL).
 FT CA_BIND 139 150 EF-HAND 4 (POTENTIAL).
 FT CONFLICT 54 54
 FT HELIX 12 25
 FT TURN 27 28
 FT STRAND 33 33
 FT HELIX 35 44
 FT TURN 45 45
 FT HELIX 51 61
 FT TURN 63 64
 FT STRAND 69 69
 FT HELIX 71 102
 FT TURN 104 105
 FT STRAND 110 110
 FT HELIX 112 121
 FT TURN 122 123
 FT HELIX 128 138
 FT STRAND 146 146
 FT HELIX 148 156
 SQ SEQUENCE 161 AA; 17546 MW; A7324A5C0B0CBDB9 CRC64;
 Query Match 0.6%; Score 7; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred. No. 87;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 487 VDIDGDG 493
 DB 138 VDIDGDG 144

RESULT 59
 YF87_METJA STANDARD; PRT; 171 AA.
 ID YF87_METJA
 AC Q58982;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical protein M1587.
 GN M1587.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcales; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geobach N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Uterback T.R., Kelley J.W., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii."
 RL Science 273:1058-1073(1996).
 CC -!- SIMILARITY: TO M.JANNASCHII MJ0417.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 CC EMBL; U67599; AAB93616.1; -
 CC PIR; B64498; B64498.
 CC TIGR; MJ1587; -
 CC InterPro; IPR003141; PHP_N.
 CC Pfam; PF02231; PHP_N; 1.
 CC SMART; SM00481; POLI1AC; 1.
 CC Hypothetical protein; Complete proteome.
 CC SEQUENCE 171 AA; 19198 MW; 7D4F7B936C5C7472 CRC64;
 KW
 QY 344 ALGDRIF 350
 DB 123 ALGDRIF 129
 Query Match 0.6%; Score 7; DB 1; Length 171;
 Best Local Similarity 100.0%; Pred. No. 91;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 344 ALGDRIF 350
 DB 123 ALGDRIF 129
 RESULT 60
 YD2F_SCHPO STANDARD; PRT; 176 AA.
 ID YD2F_SCHPO
 AC Q10263;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Very hypothetical protein C56F8.15 in chromosome I.
 GN SPAC56F8.15.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_TaxID=4896;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Horsley T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Stevers K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritzc C., Holzner E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Furnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomyces pombe.";
RL Nature 415:871-880(2002)
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; Z69728; CAA93586.1; --
CC PIR; T38925; T38925.
CC GeneDB SPombe; SPAC56F8.15; --
KW Hypothetical protein; Transmembrane.
FT TRANSMEM 3 23 POTENTIAL.
FT TRANSMEM 27 47 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 107 127 POTENTIAL.
SQ SEQUENCE 176 AA; 20595 MW; 80F48B199DA622A CRC64;

Query Match 0.6%; Score 7; DB 1; Length 176;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 820 YTL5FDT 826
DB 138 YTL5FDT 144

RESULT 61
HIS7_CHLITE STANDARD; PRT; 194 AA.
AC Q8XEF4;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19) (IGPD).
GN HISB OR CT0735.
OS Chlorobium tepidum.
OC Bacteri; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1097;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-TLS / ATCC 49652 / DSM 12025;
RX MEDLINE=22103685; PubMed=12093901;

RA Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA Dodson R.J., Deboy R., Gwinn M.L., Nelson W.C., Haft D.H.,
RA Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.L., Yang F.,
RA Holt I., Umayam L.A., Mason T., Brenner M., Shea T.P., Parksey D.,
RA Nierman W.C., Feldlyum T.V., Hansen C.L., Craven M.B., Radune D.,
RA Vamathavan J., Khouri H., White O., Gruber T.M., Ketchum K.A.,
RA Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.; TLS, a
RT "The complete genome sequence of Chlorobium tepidum.";
RT Photosynthetic, anaerobic, green-sulfur bacterium.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
CC
CC -1- CATALYTIC ACTIVITY: D-erythro-1-(imidazol-4-yl)glycerol 3-
CC phosphate = 3-(imidazol-4-yl)-2-oxopropyl phosphate + H(2)O.
CC
CC -1- PATHWAY: Histidine biosynthesis; sixth step
CC
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (by similarity)
CC
CC -1- SIMILARITY: Belongs to the imidazoleglycerol-phosphate dehydratase
CC family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC
CC EMBL; AB012844; AMW1792.1; --
CC TIGR; CT0735; --
CC HAMAP; MF_00076; --; 1.
CC InterPro; IPR000807; IGPD.
CC Pfam; PF00475; IGPD; 1.
CC ProDom; PD002282; IGPD; 1.
CC PROSITE; PS00954; IGP_DEHYDRATASE_1; 1.
CC PROSITE; PS00955; IGP_DEHYDRATASE_2; 1.
KW Histidine biosynthesis; Lyase; Complete proteome.
SQ SEQUENCE 194 AA; 21189 MW; D87295AAC9B830F2 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 194;
Best Local Similarity 100.0%; Pred. No. 1e-02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 341 IVDALGD 347
DB 80 IVDALGD 86

RESULT 62
Z239_MOUSE STANDARD; PRT; 201 AA.
ID Z239_MOUSE
AC P24399;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Zinc finger protein 239 (Zfp-239) (Zinc finger protein MOK-2).
GN ZNF239 OR ZFP239 OR MOK2 OR MOK-2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90097859; PubMed=2104662;
RA Ernault-Lange M., Kress M., Hamer D.;
RT "A gene that encodes a protein consisting solely of zinc finger
RT domains is preferentially expressed in transformed mouse cells.";
RL Mol. Cell. Biol. 10:418-421(1990).
CC
CC -1- FUNCTION: May function as a transcription factor.
CC
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).
CC
CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN TRANSFORMED
CC MOUSE CELLS.
CC
CC -1- SIMILARITY: BELONGS TO THE KRUEPPEL FAMILY OF C2H2-TYPE ZINC-
CC FINGER PROTEINS.
CC
CC -1- SIMILARITY: Contains 7 C2H2-type zinc fingers.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: M32057; AAA39729.1; --
 CC PIR: I57505; I57505.
 CC HSSP: P08047; I5P2.
 CC TRANSFAC: T00510; --
 CC MGD: MGI:1306812; Zfp239.
 CC InterPro: IPR007087; Znf_C2H2.
 CC InterPro: IPR007086; Znf_C2H2_sub.
 CC Pfam: PF00096; zf-C2H2; 7.
 CC PRINTS: PR00048; ZINC_FINGER.
 CC ProDom: PD000003; Znf_C2H2; 6.
 CC SMART: SM00355; Znf_C2H2; 7.
 CC PROSITE: PS00268; ZINC_FINGER_C2H2_1; 7.
 CC PROSITE: PS00157; ZINC_FINGER_C2H2_2; 7.
 CC Transcription regulation; DNA-binding; Zinc-finger; Metal-binding;
 CC Nuclear protein; Repeat.
 CC ZN_FING 6 28 C2H2-TYPE 1.
 CC ZN_FING 34 56 C2H2-TYPE 2.
 CC ZN_FING 62 84 C2H2-TYPE 3.
 CC ZN_FING 90 112 C2H2-TYPE 4.
 CC ZN_FING 118 140 C2H2-TYPE 5.
 CC ZN_FING 146 168 C2H2-TYPE 6.
 CC ZN_FING 174 196 C2H2-TYPE 7.
 CC SEQUENCE 201 AA; 22832 MW; B7A18COAADAB1384 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 201;
 Best Local Similarity 100.0%; Pred. No. 1e+02; 0;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 959 FTRSSSL 965
 DB 15 FTRSSSL 21

RESULT 63

GPBB_MOUSE STANDARD; PRT; 206 AA.
 AC P56400;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Platelet glycoprotein Ib beta chain precursor (GP-Ib beta) (GPtBB)
 DE (GPtB-beta).
 GN GPtBB.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97403789; PubMed=9259114;
 RA Kitaguchi T., Murata M., Ando H., Moriki T., Ikeda Y.;
 RT "Characterization of the gene encoding mouse platelet glycoprotein Ib
 RT beta.";
 RL Thromb. Res. 87:235-244(1997).
 CC -!- FUNCTION: GP-Ib, a surface membrane protein of platelets,
 CC participates in the formation of platelet plugs by binding to von
 CC Willebrand factor, which is already bound to the subendothelium
 CC (By similarity).
 CC -!- SUBUNIT: GP-Ib alpha and beta are disulfide linked. GP-IX is
 CC complexed with the GP-Ib heterodimer via a non covalent linkage
 CC (By similarity).
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -!- MISCELLANEOUS: Platelet activation apparently involves disruption
 CC of the macromolecular complex of GP-Ib with the platelet
 CC glycoprotein IX (GP-IX) and dissociation of GP-Ib from the actin-

CC binding protein.
 CC -!- SIMILARITY: Contains 1 leucine-rich (LRR) repeat.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: AB001419; BAA22424.1; --
 CC MGD: MGI:107852; Gplbb.
 CC InterPro: IPR001611; LRR.
 CC InterPro: IPR000483; LRR_Cterm.
 CC InterPro: IPR00372; LRR_Nterm.
 CC Pfam: PF00560; LRR; 1.
 CC Pfam: PF01463; LRRCT; 1.
 CC Pfam: PF01462; LRRNT; 1.
 CC SMART: SM00082; LRRCT; 1.
 CC SMART: SM00013; LRRNT; 1.
 CC Platelet; Transmembrane; Glycoprotein; Hemostasis; Blood coagulation;
 CC Signal; Phosphorylation; Cell adhesion; Leucine-rich repeat.
 CC SIGNAL 1 26 BY SIMILARITY.
 CC CHAIN 27 206 PLATELET GLYCOPROTEIN IB BETA CHAIN.
 CC DOMAIN 27 147 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 148 172 POTENTIAL.
 CC DOMAIN 173 206 CYTOPLASMIC (POTENTIAL).
 CC REPEAT 60 83 LRR.
 CC CARBOHYD 66 66 N-LINKED (GLCNAC...) (POTENTIAL).
 CC MOD RES 191 191 PHOSPHORYLATION (BY PKA) (BY SIMILARITY).
 CC SEQUENCE 206 AA; 21762 MW; AC4BCB4DFA226FID CRC64;

Query Match 0.6%; Score 7; DB 1; Length 206;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159
 DB 13 LLLALL 19

RESULT 64

PSMB_SULSO STANDARD; PRT; 208 AA.
 AC Q9UXF3;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Proteasome beta subunit precursor (EC 3.4.25.1) (Multicatalytic
 DE endopeptidase complex beta subunit).
 GN PSMB OR SSO0766 OR C40_002.
 OS Sulfolobus solfataricus.
 CC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;
 CC Sulfolobus.
 CC NCBI_TaxID=2287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=20165948; PubMed=10701121;
 RA Charlebois R.L., Singh R.K., Chan-Weiher C.C.-Y., Allard G., Chow C.,
 RA Confalonieri F., Curtis B., Duguet M., Krause G., Faguy D.,
 RA Gaucherland T., Garrett R.A., Gordon P., Jeffries A.C., Kozera C.,
 RA Kishwana N., Lailaur B., Medina N., Peng X., Penny S.L., She Q.,
 RA St Jean A., van der Oost J., Young F., Zivanovic Y., Doolittle W.F.,
 RA Ragan M.A., Sensen C.W.;
 RT "Gene content and organization of a 281-kbp contig from the genome of
 RT the extremely thermophilic archaeon, Sulfolobus solfataricus P2.";
 RL Genome 43:116-136(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX STRAIN=ATCC 35092 / DSM 1617 / P2;
 RX MEDLINE=21332296; PubMed=11427726;

RA She Q., Singh R.K., Confalonieri P., Zivanovic Y., Allard G.,
 RA Aways M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
 RA De Moors A., Brauso G., Fletcher C., Gordon P.M.K.,
 RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
 RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
 RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
 RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.,
 RA The complete genome of the crenarchaeon *Sulfolobus solfataricus* P2."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
 CC -!- FUNCTION: The proteasome is a multicatalytic proteinase complex
 CC which is characterized by its ability to cleave peptides with Arg,
 CC Phe, Tyr, Leu, and Glu adjacent to the leaving group at neutral or
 CC slightly basic pH (By similarity).
 CC -!- CATALYTIC ACTIVITY: Cleavage at peptide bonds with very broad
 CC specificity.
 CC -!- SUBUNIT: Composed of two subunits, alpha and beta. The complex is
 CC formed of four rings. The two outer rings are each composed of
 CC seven alpha subunits. The two inner rings are each composed of
 CC seven beta subunits (By similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- SIMILARITY: Belongs to peptidase family T1B.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: Y18930; CAB57537.1; ALT_INIT.
 CC EMBL: AE008700; AAK41063.1; ALT_INIT.
 CC HSP: P28061; 1PXA.
 CC MEROPS: T01.002; -.
 CC InterPro: IPR000243; Pept T1A sub.
 CC Pfam: PF001353; Peptidase_T1.
 CC Pfam: PF00227; proteasome; 1.
 CC PRINTS: PR00141; PROTEASOME.
 CC PROSITE: PS00854; PROTEASOME_B; 1.
 CC Proteasome; Hydrolase; Protease; Complete proteome;
 KW Threonine protease.
 KW
 FT PROPEP 1 14 APPARENTLY REMOVED IN MATURE FORM (BY
 FT SIMILARITY).
 FT CHAIN 15 208 PROTEASOME BETA SUBUNIT.
 FT ACT_SITE 15 15 PROBABLE.
 FT SEQUENCE 208 AA; 22927 MW; C940A8273AB25CB9 CRC64;
 Query Match 0.6%; Score 7; DB 1; Length 208;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 686 TTTVGIR 692
 Db 15 TTTVGIR 21
 RESULT 65
 KPTA PYRAE STANDARD; PRT; 213 AA.
 AC Q823P2;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Probable RNA 2'-phosphotransferase (EC 2.7.-.-).
 GN KPTA OR PAR3647.
 OS Pyrobaculum aerophilum.
 OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
 OC Crenarchaeota; Pyrobaculum.
 OC NCBI_TaxID=13773;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=IM2 / ATCC 51768 / DSM 7523;
 CC MEDLINE=21664397; PubMed=11792869;

RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
 RA Miller J.H.;
 RT "Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum*
 RT aerophilum";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
 CC -!- FUNCTION: Removes the 2'-phosphate from RNA via an intermediate in
 CC which the phosphate is ADP-ribosylated by NAD followed by a
 CC presumed transesterification to release the RNA and generate ADP-
 CC ribose 1''-2''-cyclic phosphate (APPR>P). May function as an ADP-
 CC ribosylase (By similarity).
 CC -!- SIMILARITY: Belongs to the kpta / Tpt1 family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: AE009943; AAL65071.1; -.
 CC HAMAP: MF_00399; -; 1.
 CC InterPro: IPR002745; Ptrans_Kpta/Tpt1.
 CC Pfam: PF01885; PFS_2-RNA; 1.
 CC ProDom: PD009619; P04transfrase; 1.
 CC Transfrase; NAD; Complete proteome.
 KW SEQUENCE 213 AA; 23452 MW; 12DF2AD12CB69636 CRC64;
 SQ
 Query Match 0.6%; Score 7; DB 1; Length 213;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 72 DVYKCPV 78
 Db 3 DVYKCPV 9
 RESULT 66
 CH13 HUMAN STANDARD; PRT; 214 AA.
 ID CH13 HUMAN
 AC Q96KS9; Q9N3M3; Q9NSR0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein C8orf13.
 GN C8orf13.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Appel S., Bergheim A., Reichwald K., Reis A., Rosenthal A., Ramsay M.,
 RA Hennies H.;
 RT "Transcript map of the KWE critical region on chromosome 8p22-p23."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=amygdala;
 RA Wambutt R., Heubner D., Mewes H.-W., Weil B., Wiemann S.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE OF 83-214 FROM N.A.
 RC TISSUE=amygdala;
 RA Ottenwaelder B., Obermaier B., Mewes H.-W., Weil B., Wiemann S.;
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SEC FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALL 1159
 Db 6 LLLALL 12

RESULT 69
 SDFL HUMAN STANDARD; PRT; 221 AA.
 AC O9CNG; O9BRIS;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Stromal cell-derived factor 2-like protein 1 precursor (SDF2 like
 protein 1) (PwP1-interacting protein 8).
 GN SDF2L1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=21092619; PubMed=11162531;
 RA Fukuda S., Sumii M., Masuda Y., Takahashi M., Koike N., Teishima J.,
 Yasumoto H., Itamoto T., Asahara T., Dohi K., Kamiya K.,
 RT "Murine and human SDF2L1 is an endoplasmic reticulum stress-inducible
 RT gene and encodes a new member of the Pnt/rt protein family.";
 RL Biochem. Biophys. Res. Commun. 280:407-414(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Uterus;
 RA Honore B.;
 RT "hPwP1-interacting protein 8.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=4238257; PubMed=12477932;
 RA Klausner R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uscin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Fahey J., Helton E., Kettner M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.N., Small D.E.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalski U., Marra M.A.;
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (Potential).
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed with high expression in
 CC testis, moderate expression in the pancreas, spleen, prostate,
 CC small intestine and colon. Very low expression is seen in brain
 CC and skeletal muscle.
 CC -!- SIMILARITY: Contains 3 MIR domains.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

or send an email to license@isb-sib.ch).

EMBL; AB043007; BAB18277.1; -
 DR EMBL; AF277316; AAK69113.1; -
 DR EMBL; BC006248; AAK06248.1; -
 DR PIR; JC7587; JC7587.
 DR Genew; HGNC:10676; SDF2L1.
 DR MIN; 607551; -
 DR InterPro; IPR000886; ER target_S.
 DR InterPro; IPR003608; MIR.
 DR Pfam; PF02815; MIR; 3.
 DR SMART; SMO0472; MIR; 3.
 DR PROSITE; PS00014; ER TARGET; 1.
 DR PROSITE; PS0919; MIR; 3.
 KW Endoplasmic reticulum; Signal; Repeat.
 FT SIGNAL 1 28
 FT CHAIN 29 221
 FT STROMAL CELL-DERIVED FACTOR 2-LIKE
 FT PROTEIN 1.
 FT DOMAIN 33 87
 FT MIR 1.
 FT DOMAIN 95 150
 FT MIR 2.
 FT DOMAIN 151 205
 FT MIR 3.
 FT SITE 218 221
 FT PREVENT SECRETION FROM ER (POTENTIAL).
 FT C -> R (IN REF. 2 AND 3).
 FT CONFLICT 94 94
 FT CONFLICT 162 162
 FT L -> F (IN REF. 2 AND 3).
 FT CONFLICT 162 162
 FT SEQUENCE 221 AA; 23511 MW; 23EBE14BE92EF CRC64;
 SQ

Query Match 0.6%; Score 7; DB 1; Length 221;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1160
 Db 17 LLLALLV 23

RESULT 70
 EMBP_MOUSE
 ID EMBP_MOUSE STANDARD; PRT; 223 AA.
 AC Q61878;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Eosinophil granule major basic protein precursor (MBP) (Proteoglycan
 DE 2, bone marrow).
 DE PRG2 OR MBP-1.
 GN Mus musculus (Mouse).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RA STRAIN=CBA/J; TISSUE=Liver;
 RX MEDLINE=95403975; PubMed=7673718;
 RA Larson K.A., Horton M.A., Madden B.J., Gleich G.J., Lee N.A.,
 RA Lee J.J.;
 RT "The identification and cloning of a murine major basic protein gene
 RT expressed in eosinophils.";
 RL J. Immunol. 155:3002-3012(1995).
 CC -!- FUNCTION: Cytotoxin and helminthotoxin. MBP also induces
 CC noncytolytic histamine release from basophils. It is involved in
 CC antiparasitic defense mechanisms and immune hypersensitivity
 CC reactions (By similarity).
 CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
 CC granule (crystalloid core) (By similarity).
 CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

CC EMBL; L46768; AAA83027.1; -
DR HSP; P13727; IHRU.
DR MGI; MGI:103294; Prg2.
DR InterPro; IPR002352; Emaior_bas.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR00770; EmaiorBASICP.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; FALSE_NEG.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin.
FT SIGNAL 1 16
FT PROPEP 17 106
FT CHAIN 107 223
FT DOMAIN 124 223
FT DISULFID 126 221
FT DISULFID 198 213
SQ SEQUENCE 223 AA; 24255 MW; 7D6D946DCEADA00 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1160
DB 5 LLLALLV 11

RESULT 71
EMBP RAT STANDARD; PRT; 227 AA.
AC Q63189;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein precursor (MBP).
GN PRG2 OR MBP
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE FROM N.A.
RA STRAIN=Sprague-Dawley; TISSUE=Bone marrow;
RA MEDLINE=56138543; PubMed=8547309;
RA Nitton T., Watanabe M., Okayama H., Misawa S., Isobe Y., Hayashi H.,
RA Mue S., Ohuchi K.;
RA "Cloning of cDNA for rat eosinophil major basic protein.";
RL Biochim. Biophys. Acta 1264:261-264(1995).
CC -!- FUNCTION: Cytochrome and helminthotoxin. MBP also induces
CC nontoxic histamine release from basophils. It is involved in
CC antiparasitic defense mechanisms and immune hypersensitivity
CC reactions (by similarity).
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule (crystalloid core) (By similarity).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D50568; BAA09129.1; -
DR PIR; S68150; S68150.
DR HSP; P13727; IHRU.
DR InterPro; IPR002352; Emaior_bas.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR00770; EmaiorBASICP.

```

```

DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin.
FT SIGNAL 1 16
FT PROPEP 17 110
FT CHAIN 111 227
FT DOMAIN 128 227
FT DISULFID 130 225
FT DISULFID 202 217
FT DOMAIN 59 68
SQ SEQUENCE 227 AA; 25129 MW; 34F2B5514090DE82 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 227;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1160
DB 5 LLLALLV 11

RESULT 72
EMBI_CAVPO STANDARD; PRT; 233 AA.
ID EMBI_CAVPO
AC P22032;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Eosinophil granule major basic protein 1 precursor (MBP-1).
GN MBP1.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
SEQUENCE FROM N.A., AND SEQUENCE OF 115-162.
RC TISSUE=Eosinophil;
RX MEDLINE=91160746; PubMed=1705901;
RA Aoki I., Shindoh Y., Nishida T., Nakai S., Hong Y.-M., Mio M.,
RA Saito T., Tasaka K.;
RA "Sequencing and cloning of the cDNA of guinea pig eosinophil major
RA basic protein.";
RL FEBS Lett. 279:330-334(1991).
RN [2]
PARTIAL SEQUENCE.
RX MEDLINE=94092714; PubMed=8268206;
RA Hashimoto Y., Nagaoka I., Yamashita T.;
RA "Purification of the antibacterial fragments of guinea-pig major
RA basic protein.";
RL Biochim. Biophys. Acta 1203:236-242(1993).
CC -!- FUNCTION: MBP may play some important roles in the allergic
CC reactions and inflammations, since MBP is capable of releasing
CC histamine from mast cells and damaging the epithelial cells of
CC bronchial tubes. Antiparasitic and antibiotic.
CC -!- SUBCELLULAR LOCATION: Matrix of eosinophil's large specific
CC granule (crystalloid core).
CC -!- SIMILARITY: Contains 1 C-type lectin family domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D90251; BAA14291.1; -
DR PIR; S13625; S13625.
DR HSP; P13727; IHRU.
DR InterPro; IPR002352; Emaior_bas.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00059; lectin c; 1.
DR PRINTS; PR00059; lectin c; 1.

```

```
DR PRINTS; PRO0770; EMAJORBASICP.
DR SMART; SM00034; CLECT.1.
DR PROSITE; PS00615; C-TYPE LECTIN 1; 1.
DR PROSITE; PS00641; C-TYPE LECTIN 2; 1.
KW Eosinophil; Signal; Immune response; Antibiotic; Lectin;
KW Multigene family.
FT SIGNAL 1 15 POTENTIAL.
FT PROPEP 16 114 ACIDIC.
FT CHAIN 115 233 EOSINOPHIL GRANULE MAJOR BASIC PROTEIN 1.
FT DOMAIN 132 233 C-TYPE LECTIN (SHORT FORM).
FT DISULFID 134 231 BY SIMILARITY.
FT DISULFID 208 223 BY SIMILARITY.
SQ SEQUENCE 233 AA; 26268 MW; C8D5E96D927C55C8 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 1153 LLLALL 1159
DB 4 LLLALL 10

RESULT 73
YBEU ECOLI
ID YBEU ECOLI STANDARD; PRT; 235 AA.
AC P77427;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ybeU.
GN YBEU OR B0648.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC SEQUENCE FROM N.A.
DR EMBL; U28377; AAA69134.1; -.
DR EMBL; AF000379; AAC76003.1; -.
DR PIR; E65082; E65082.
DR EcoGene; EG12987; YGSA.
DR InterPro; IPR007563; DUF554.
DR Pfam; PF04474; DUF554; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
SQ SEQUENCE 235 AA; 24614 MW; CC89D2A93FED29EA CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;
```

```
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE000169; AAC73749.1; -.
CC EMBL; U82598; AAB40849.1; -.
DR EMBL; D90704; BAA35295.1; -.
DR EMBL; D90705; BAA35300.1; -.
DR PIR; F64799; F64799.
DR EcoGene; EG13651; ybeU.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 235 AA; 27022 MW; 9A1CAF5D6A0ECF1 CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;

QY 926 ELAAGSD 932
DB 202 ELAAGSD 208

RESULT 74
YOGA ECOLI
ID YOGA ECOLI STANDARD; PRT; 235 AA.
AC Q46831;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ygaA.
GN YOGA OR B2966.
OS Escherichia coli.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC SEQUENCE FROM N.A.
DR EMBL; U28377; AAA69134.1; -.
DR EMBL; AF000379; AAC76003.1; -.
DR PIR; E65082; E65082.
DR EcoGene; EG12987; YGSA.
DR InterPro; IPR007563; DUF554.
DR Pfam; PF04474; DUF554; 1.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 2 22 POTENTIAL.
FT TRANSMEM 34 54 POTENTIAL.
FT TRANSMEM 56 76 POTENTIAL.
FT TRANSMEM 102 122 POTENTIAL.
FT TRANSMEM 147 167 POTENTIAL.
FT TRANSMEM 178 198 POTENTIAL.
FT TRANSMEM 210 230 POTENTIAL.
SQ SEQUENCE 235 AA; 24614 MW; CC89D2A93FED29EA CRC64;

Query Match 0.6%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02; Indels 0;
Matches 7; Conservative 0; Mismatches 0; Gaps 0;
```

```

OY 1151 GGLLLA 1157
DR 194 GGLLLA 200
DB

RESULT 75
RR2_MAIZE
ID RR2_MAIZE STANDARD; PRT; 236 AA.
AC P16037;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Chloroplast 30S ribosomal protein S2.
GN RPS2.
OS Zea mays (Maize).
OC Chloroplast.
OC Eukaryota, Viridiplantae, Streptophyta, Embryophyta, Tracheophyta,
OC Spermatophyta, Magnoliophyta, Liliopsida, Poales, Poaceae;
OC PACAD clade, Panicoideae, Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RN SEQUENCE FROM N.A.
RX MEDLINE=90175001; PubMed=2308853;
RA Igloi G.L., Meinke A., Doery I., Koessel H.;
RT "Nucleotide and derived amino acid sequence of rps2 from maize
RT chloroplasts.";
RL Nucleic Acids Res. 18:663-663(1990).
RN [2]
RN SEQUENCE FROM N.A.
RX STRAIN=CV_F99CMSSR37; TISSUE=Leaf;
RX MEDLINE=90272437; PubMed=2140888;
RA Stahl D., Rodermeil S., Subramanian A.R., Bogorad L.;
RT "Nucleotide sequence of a 3.46 kb region of maize chloroplast DNA
RT containing the gene cluster rpoC2-rps2-atpI-atpH.";
RL Nucleic Acids Res. 18:3073-3074(1990).
RN [3]
RN SEQUENCE FROM N.A.
RX MEDLINE=95395841; PubMed=7666415;
RA Maier R.M., Neckermann K., Igloi G.L., Koessel H.;
RT "Complete sequence of the maize chloroplast genome: gene content,
RT hotspots of divergence and fine tuning of genetic information by
RT transcript editing.";
RL J. Mol. Biol. 251:614-628(1995).
RN [4]
RN SEQUENCE OF 1-36 FROM N.A.
RX MEDLINE=90340289; PubMed=2381419;
RA Igloi G.L., Meinke A., Doery I., Koessel H.;
RT "Nucleotide sequence of the maize chloroplast rpo B/C1/C2 operon:
RT comparison between the derived protein primary structures from
RT various organisms with respect to functional domains.";
RL Mol. Gen. Genet. 221:379-394(1990).
CC -!- SIMILARITY: Belongs to the S2P family of ribosomal proteins.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X17318; CAA35198.1; -
CC EMBL; X52270; CAA36512.1; -
CC EMBL; X86563; CAA60273.1; -
CC PIR; S08249; R32M2.
CC Gramene; P16037; -
CC MaizEDB; 66014; -
CC HAMAP; MF_00291; -; 1.
CC InterPro; IPR001865; Ribosomal_S2.
CC Pfam; PF00318; Ribosomal_S2; 1.
CC PRINTS; PR00395; RIBOSOMALS2.
CC TIGRFAMs; TIGR01011; rpsB_Bact; 1.

```

```

DR PROSITE; PS00962; RIBOSOMAL_S2_1; 1.
DR PROSITE; PS00963; RIBOSOMAL_S2_2; 1.
KW Ribosomal protein; Chloroplast.
SQ SEQUENCE 236 AA; 26881 MW; 80A3F99A8955SFAD5 CRC64;

Query Match
Best Local Similarity 100.0%; Score 7; DB 1; Length 236;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 DIVIVLD 170
DB 163 DIVIVLD 169

Search completed: September 21, 2004, 13:32:10
Job time : 34 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:22:25 ; Search time 74 Seconds
(without alignments)
5065.346 Million cell updates/sec

Title: US-09-980-403-2
Perfect score: 1188
Sequence: 1 MDLPRLVAVWALSMPGFT.....PSARRRRPGLDTPKVL 1188

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1017041 seqs, 315518202 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : SPTREMBL 25:
1: sp archaea.*
2: sp bacteria.*
3: sp fungi.*
4: sp human.*
5: sp invertebrate.*
6: sp mammal.*
7: sp mhc.*
8: sp organelle.*
9: sp phage.*
10: sp plant.*
11: sp rodent.*
12: sp virus.*
13: sp vertebrate.*
14: sp unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	662	55.7	823	4	Q8WY18
2	122	10.3	1188	11	Q7TQC3
3	45	3.8	1823	11	Q8C884
4	15	1.3	1171	13	Q42094
5	11	0.9	288	11	Q8EM12
6	10	0.8	116	13	Q91670
7	10	0.8	399	16	Q8XSH5
8	9	0.8	164	4	Q8N112
9	9	0.8	315	16	Q8YCE3
10	9	0.8	316	16	Q8FVL4
11	9	0.8	400	6	Q8HY16
12	9	0.8	413	16	Q89LS8
13	9	0.8	497	13	Q73804
14	9	0.8	607	10	Q39775
15	9	0.8	780	13	Q06271
16	8	0.7	54	6	Q7TS65

90 Q8pg60 xanthomonas
 91 Q9p3u1 schizosacch
 92 Q8ye57 brucella me
 93 Q8fyx8 brucella su
 94 Q04611 arabidopsis
 95 Q8ewh6 mycoplasma
 96 Q9ey90 amycolatops
 97 Q821z0 streptomyce
 98 Q52642 pseudomonas
 99 Q51991 pseudomonas
 100 Q64584 rattus norv

ALIGNMENTS

RESULT 1
 Q8WY18
 ID Q8WY18 PRELIMINARY; PRT; 823 AA.
 AC Q8WY18
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-JUN-2003 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE MSTP018.
 GN MST018.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 CX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Aorta;
 RA Xu Y.Y., Sun L.Z., Wu Q.Y., Liu Y.Q., Liu B., Zhao B., Wang X.Y.,
 RA Song L., Ye J., Sheng H., Gao Y., Zhang C.L., Zhang J., Wei Y.J.,
 RA Sun Y.H., Jiang Y.X., Zhao X.W., Liu L.S., Ding J.F.,
 RA Gao R.L., Qiang B.Q., Yuan J.G., Liew C.C., Zhao M.S., Hui R.T.;
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF111799; AAL39001.1;
 DR GO; GO:0008305; C:integrin complex; IEA.
 DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
 DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
 DR InterPro; IPR000413; Integrin_alpha.
 DR Pfam; PF01839; FG-GAP; 3.
 DR PRINTS; PR01185; INTEGRINA.
 DR SMART; SM00191; Int alpha; 4.
 SQ SEQUENCE 823 AA; 92672 MW; DE4E78079DCD4925 CRC64;
 Query Match 55.7%; Score 662; DB 4; Length 823;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 762; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 366 MSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKPPPELKNHGAYL 425
 DB 1 MSQTGFSSHHVDEGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKPPPELKNHGAYL 60
 QY 426 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLLTHQAMRQGIQGSYFGSBIT 485
 DB 61 GYTVTSVSSRQGRVYVAGAPRNFHTGKVLFTMHNRSLLTHQAMRQGIQGSYFGSBIT 120
 QY 486 SVDIDGDGVTDLVLCAPMYNEGERGKVYVELRQNFVYNGTLKDSHSYQNAFQSS 545
 DB 121 SVDIDGDGVTDLVLCAPMYNEGERGKVYVELRQNFVYNGTLKDSHSYQNAFQSS 180
 QY 546 IASVRDLNDSYNDVVVGAPLEDNHNAGAIYIFHGFSGILTKPKORITASELATGLQYFG 605
 DB 181 IASVRDLNDSYNDVVVGAPLEDNHNAGAIYIFHGFSGILTKPKORITASELATGLQYFG 240
 QY 605 CSTHGOLDNEGLDLAVGALGNVILWSRPVQINASLHFPSPKINIFHRDCKRSGRD 665
 DB 241 CSTHGOLDNEGLDLAVGALGNVILWSRPVQINASLHFPSPKINIFHRDCKRSGRD 300
 QY 666 ATCLAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTTPRAHLDGGDRFTNRAVLLSSG 725

DB 301 ATCLAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTTPRAHLDGGDRFTNRAVLLSSG 360
 QY 726 QELCERINPHVLDIADYKVPVTFVSVEYSLEDPDHGMDDGWPPTLRVSVFPWNGCNEDE 785
 DB 361 QELCERINPHVLDIADYKVPVTFVSVEYSLEDPDHGMDDGWPPTLRVSVFPWNGCNEDE 420
 QY 786 HCVPLDLVLDASDLPTAMEYQQRVLRKPAQCSAYTISFDITVFIESTRORVAEATLE 845
 DB 421 HCVPLDLVLDASDLPTAMEYQQRVLRKPAQCSAYTISFDITVFIESTRORVAEATLE 480
 QY 846 NRGENAYSTVLNISQSANLQFASLIQKEDSDSGSIECVNEERLQKQVCNVPYFFRAKAK 905
 DB 481 NRGENAYSTVLNISQSANLQFASLIQKEDSDSGSIECVNEERLQKQVCNVPYFFRAKAK 540
 QY 906 VAFRLDSFSSKIFLHLEIELAAGSDSNEDSKEDNAPLRPHLYKYEADVLETRSSSL 965
 DB 541 VAFRLDSFSSKIFLHLEIELAAGSDSNEDSKEDNAPLRPHLYKYEADVLETRSSSL 600
 QY 966 SHYEVKUNSLERYDGIGPPFSCIFRIQNLGLFPIHGMMKITIPIATRSNRLKLKRD 1025
 DB 601 SHYEVKUNSLERYDGIGPPFSCIFRIQNLGLFPIHGMMKITIPIATRSNRLKLKRD 660
 QY 1026 LTDEANTSCNIGWNSTEYRPTVEEDLRAPQLNHSNDVVSINCNILVYVQNEINPHLL 1085
 DB 661 LTDEANTSCNIGWNSTEYRPTVEEDLRAPQLNHSNDVVSINCNILVYVQNEINPHLL 720
 QY 1086 GNLWLRSLKALKYKSMKIMVNAALQROFHSPFIFREEDPSRQI 1128
 DB 721 GNLWLRSLKALKYKSMKIMVNAALQROFHSPFIFREEDPSRQI 763
 RESULT 2
 Q7TQC3
 ID Q7TQC3 PRELIMINARY; PRT; 1188 AA.
 AC Q7TQC3
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE All integrin.
 GN ITGAL1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tiger C.-F., Popova S.N., Velling T., Ekman S., Forsberg E.,
 RA Gullberg D.;
 RA "allb1 integrin is important for mesenchymal cell function:
 RT elimination of allb1 leads to dwarfism."
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Johansson M., Popova S.N.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY124460; AAM62130.1;
 KW Integrin.
 SQ SEQUENCE 1188 AA; 133067 MW; 69B2BF20D75B4271 CRC64;
 Query Match 10.3%; Score 122; DB 11; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 8.3e-118;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 81 GNTCKNLGRLVTLNVSERKDNRLGLSLATPKNSFLACSPKWSHCEGSSVYTTGMS 140
 DB 81 GNTCKNLGRLVTLNVSERKDNRLGLSLATPKNSFLACSPKWSHCEGSSVYTTGMS 140
 QY 141 RVNSNFRSKTAPALQRCQTYMDIVILVLDGNSIYPVVEQHFILNLIKFYIGPGQIQ 200
 DB 141 RVNSNFRSKTAPALQRCQTYMDIVILVLDGNSIYPVVEQHFILNLIKFYIGPGQIQ 200
 QY 201 VG 202


```
Db 201 VG 202

RESULT 3
Q8CEB4 PRELIMINARY; PRT; 823 AA.
ID Q8CEB4
AC Q8CEB4;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE MST018 homolog.
GN 4732459H24Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_TaxID=10090;
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Skin;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK028821; BAC26137.1; -.
DR MGD; MGI:2442114; 4732459H24Rik.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR PRINTS; PRO1185; INTEGRINA.
DR SMART; SM00191; Int_alpha; 4.
DR SMART; SM00191; Int_alpha; 4.
SQ SEQUENCE 823 AA; 92264 MW; A330236324A0E089 CRC64;

Query Match 3.8%; Score 45; DB 11; Length 823;
Best Local Similarity 100.0%; Pred.No. 2.5e-37; Indels 0; Gaps 0;
Matches 45; Conservative 0; Mismatches 0;

QY 766 GWPTTLRVSPFVWNGCNEDEHCVDELVDARSDLPTAMEYQORVL 810
DB 401 GWPTTLRVSPFVWNGCNEDEHCVDELVDARSDLPTAMEYQORVL 445

RESULT 4
O42094 PRELIMINARY; PRT; 1171 AA.
ID O42094
AC O42094;
DT 01-JAN-1998 (T-EMBlrel. 05, Created)
DT 01-JAN-1998 (T-EMBlrel. 05, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE ALPHAI integrin.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]_TaxID=9031;
RP SEQUENCE FROM N.A.
RC TISSUE=Gizzard;
RX MEDLINE=97476270; PubMed=9334246;
RA Obata H., Hayashi K., Nishida W., Momiyama T., Uchida A., Ochi T.,
RA Sobue K.;
RT "Smooth muscle cell phenotype-dependent transcriptional regulation of
RT the alphaI integrin gene.";
RL J. Biol. Chem. 272:26643-26651 (1997).
DR EMBL; AB000470; BAA23160.1; -.
DR EMBL; AB000471; BAA23161.1; -.
DR PIR; A55348; A55348.
DR HSSP; P17301; 1A0X.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
```

```

DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Integrin alpha-1 (Fragment).
GN GENE 4.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312514; PubMed=8700860;
RA Brown D.D., Wang Z., Furlow J.D., Kanamori A., Schwartzman R.A.,
RA Reno B.F., Pinder A.;
RT "The thyroid hormone-induced tail resorption program during Xenopus
RT laevis metamorphosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:1924-1929 (1996).
DR EMBL; U4025; AAC59878.1; -
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR SMART; SMO0191; Int_alpha; 1.
FT NON_TER 116
SQ SEQUENCE 116 AA; 12749 MW; 2A8A5CA3E83DD07E CRC64;

Query Match 0.8%; Score 10; DB 13; Length 116;
Best Local Similarity 100.0%; Pred. No. 0.2; Mismatches 0; Indels 0; Gaps 0;
Matches 10; Conservative 0;

Qy 39 AFGGYTVQOH 48
Db 39 AFGGYTVQOH 48

RESULT 7
Q8XSH5 PRELIMINARY; PRT; 399 AA.
AC Q8XSH5
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 25, Last annotation update)
DE Putative transport transmembrane protein.
GN RFP0499 OR R500378.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OG Plasmid megaplasmid.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Ralstonia.
OX NCBI_TaxID=305;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=GM11000;
RA MEDLINE=21681879; PubMed=11823852;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Bottier P., Camus J.C., Catolico L.,
RA Chandler M., Choine N., Claudel-Renard C., Cunnac S., Denange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
RA Siguer P., Thebault P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RL Nature 415:497-502 (2002).
DR EMBL; AL646079; CAD17650.1; -
DR GO; GO:0046821; C:extrachromosomal DNA; IEA.
DR GO; GO:0008457; P:protein folding; IEA.
DR InterPro; IPR001179; FKBP_PPase.
DR InterPro; IPR007114; MFS.
DR PROSITE; PS00453; FKBP_PPase_1; 1.
DR PROSITE; PS00850; MFS; 1.
KW Plasmid; Complete proteome.
SQ SEQUENCE 399 AA; 42148 MW; 096BB97E83DCA7D1 CRC64;

Query Match 0.8%; Score 10; DB 16; Length 399;
Best Local Similarity 100.0%; Pred. No. 0.59;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1152 GLLLLALLVL 1161
Db 161 GLLLLALLVL 170

RESULT 8
Q8N112 PRELIMINARY; PRT; 164 AA.
AC Q8N112
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-WAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical gene LOC132228 (Hypothetical protein FLJ38608).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Brain;
RA Strausberg R.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Nishi T., Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H.,
RA Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
RA Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
RA Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
RA Kikuchi H., Kanda K., Wagatsuma M., Murakawa K., Kanehori K.,
RA Takahashi-Fujii A., Oshima A., Sugiyama A., Kawakami B., Suzuki Y.,
RA Sugano S., Negahari K., Masuno Y., Nagai K., Isogai T.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC028000; AAH28000.1; -
DR EMBL; AK095927; BAC04652.1; -
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 17868 MW; 509993458714252E CRC64;

Query Match 0.8%; Score 9; DB 4; Length 164;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALLVL 1161
Db 99 LLLALLVL 107

RESULT 9
Q8YCR3 PRELIMINARY; PRT; 315 AA.
AC Q8YCR3
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE PERMEASE.
GN BMEII0465.
OS Brucella melitensis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29459;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=16M / ATCC 23456 / Biotype 1;
RA MEDLINE=20020109; PubMed=11756688;
RA DelVecchio V.G., Kapatral V., Redkar R.J., Patra G., Muejer C., Los T.,
RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
RA Jablonski L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
RA Selkov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-J.,
RA Haselkorn R., Kyrides N., Overbeek R.;
RT "The genome sequence of the facultative intracellular pathogen
RT Brucella melitensis.";

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448 (2002).
DR EMBL; AE009683; JAL53707.1; -.
DR PIR; AH3567; AH3567.
DR InterPro; IPR002781; DUF81.
DR Pfam; PF01925; DUF81; 1.
KW Complete proteome.
SQ SEQUENCE 315 AA; 33486 MW; 7A34836134FB2F7D CRC64;

Query Match 0.8%; Score 9; DB 16; Length 315;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162
DB 283 LLLALLVLA 291

RESULT 10
Q8FVLA PRELIMINARY; PRT; 316 AA.
ID Q8FVLA;
AC Q8FVLA;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Membrane protein, putative.
GN BRA0823.
OS Brucella suis.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Brucellaceae; Brucella.
OX NCBI_TaxID=29461;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1330 / Biovar 1;
RX MEDLINE=22247741; PubMed=12271122;
RA Paulsen I.T., Seshadri R., Nelson K.E., Eisen J.A., Heidelberg J.F.,
RA Read T.D., Dodson R.J., Umayam L., Brinkac L.M., Beanan M.J.,
RA Daugherty S.C., Deboy R.T., Durkin A.S., Kolonay J.F., Madupu R.,
RA Nelson W.C., Ayodeji B., Krahl M., Shetty J., Malek J., Van Aken S.E.,
RA Riedmiller S., Tettelin H., Gill S.R., White O., Salzberg S.L.,
RA Hoover D.L., Lindler L.E., Halling S.M., Boyle S.M., Fraser C.M.,
RT "The Brucella suis genome reveals fundamental similarities between
RT animal and plant pathogens and symbionts."
RT Proc. Natl. Acad. Sci. U.S.A. 99:13148-13153 (2002).
RL EMBL; AE014576; AAN3398.1; -.
DR TTGR; BRA0823; -.
DR InterPro; IPR002781; DUF81.
DR Pfam; PF01925; DUF81; 1.
KW Complete proteome.
SQ SEQUENCE 316 AA; 33671 MW; 4DC0DC32D8AB0807 CRC64;

Query Match 0.8%; Score 9; DB 16; Length 316;
Best Local Similarity 100.0%; Pred. No. 5.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162
DB 284 LLLALLVLA 292

RESULT 11
Q8HY16 PRELIMINARY; PRT; 400 AA.
ID Q8HY16;
AC Q8HY16;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Poliovirus receptor.
GN FVR.
OS Cebus apella (Brown-capped capuchin).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Cebinae; Cebus.
OX NCBI_TaxID=9515;
RN [1]

```

```

SEQUENCE FROM N.A.
RA Ida-Hosonuma M., Sasaki Y., Toyoda H., Nomoto A., Gotoh O.,
RA Yonetaka H., Koike S.
RT "Host range of poliovirus is restricted to simians because of a rapid
RT sequence change of the poliovirus receptor gene during evolution."
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB086131; BAC41711.1; -.
DR EMBL; AB086124; BAC41711.1; JOINED.
DR EMBL; AB086125; BAC41711.1; JOINED.
DR EMBL; AB086126; BAC41711.1; JOINED.
DR EMBL; AB086127; BAC41711.1; JOINED.
DR EMBL; AB086128; BAC41711.1; JOINED.
DR EMBL; AB086129; BAC41711.1; JOINED.
DR EMBL; AB086130; BAC41711.1; JOINED.
DR GO; GO:0004872; F:receptor activity; IEA.
DR InterPro; IPR001064; Crystallin.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF00047; Ig; 3.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS00225; CRYSTALLIN_BETAGAMMA; 1.
DR PROSITE; PS50835; IG_LIKE; 3.
KW Receptor.
SQ SEQUENCE 400 AA; 43322 MW; C6C3A59A0A285F3A CRC64;

Query Match 0.8%; Score 9; DB 6; Length 400;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVLA 1161
DB 9 LLLALLVLA 17

RESULT 12
Q89LS8 PRELIMINARY; PRT; 413 AA.
ID Q89LS8;
AC Q89LS8;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE BLR4465 protein.
GN BLR4465.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idegawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsumoto M., Shimpo S., Tsuruoka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110."
RL DNA Res. 9:189-197 (2002).
DR EMBL; AP005951; BAC49730.1; -.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0016811; F:hydrolase activity, acting on carbon-nitrog...; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR000345; CytC heme BS.
DR InterPro; IPR004304; FmdA_AmdA.
DR Pfam; PF03669; FmdA_AmdA_1.
DR PROSITE; PS00190; CYTCCHROME_C; 1.
KW Complete proteome.
SQ SEQUENCE 413 AA; 45159 MW; E2F309DCC2E9D82B CRC64;

Query Match 0.8%; Score 9; DB 16; Length 413;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Query Match 0.8%; Score 9; DB 10; Length 607;
Best Local Similarity 100.0%; Pred.No.9.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 148 FSKTVAPAL 156
DB 18 FSKTVAPAL 26

RESULT 13

ID O73804 PRELIMINARY; PRT; 497 AA.

AC O73804;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Takifugu.
OX NCBI_TaxID=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99177347; PubMed=10077531;
RA Gellner K., Brenner S.;
RT "Analysis of 148 kb of genomic DNA around the wnt1 locus of Fugu rubripes";
RL Genome Res. 9:251-258 (1999).
DR EMBL; AF056116; AAC34385.1; -;
DR InterPro; IPR008075; Lipocalin1_recep.
DR InterPro; IPR006876; LMBRI.
DR Pfam; PF04791; LMBRI; 1.
DR PRINTS; PR01692; LIPOCALINIMR.
KW Hypothetical protein.
SQ SEQUENCE 497 AA; 56376 MW; 802743B983E83E64 CRC64;

Query Match 0.8%; Score 9; DB 13; Length 497;
Best Local Similarity 100.0%; Pred.No.8;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1161
DB 169 LLLALLVL 177

RESULT 14

ID Q39775 PRELIMINARY; PRT; 607 AA.

AC Q39775;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Legumin, 11S globulin.
OS Gnetum gnemon (Bago).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Gnetales; Gnetopsida; Gnetales; Gnetales; Gnetales.
OX NCBI_TaxID=3382;
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE=Seed.
RA Shutoy A.D., Braun H., Chesnokov Y.V., Horstmann C., Kakhovskaya I.A.,
RA Baumlein H.;
RT "sequence peculiarity of Gnetalean legumin-like seed storage proteins";
RL J. Mol. Evol. 47:486-492 (1998).
DR EMBL; Z50779; CAA90642.1; -;
DR PIR; S60658; S60658.
DR GO; GO:0045733; r:nutrient reservoir activity; IEA.
DR InterPro; IPR006045; Cupin.
DR InterPro; IPR006044; Seedstore_11s.
DR Pfam; PF00190; Cupin; 2.
DR PRINTS; PR00439; 11SGLOBULIN.
SQ SEQUENCE 607 AA; 68910 MW; 7AD6DDC604FB3A60 CRC64;

Query Match 0.8%; Score 9; DB 10; Length 607;
Best Local Similarity 100.0%; Pred.No.9.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLVL 1161
DB 10 LLLALLVL 18

RESULT 15

ID Q06271 PRELIMINARY; PRT; 780 AA.

AC Q06271;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin alpha-2 subunit (Fragment).
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RA Meng F., Desimone D.W.;
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RX MEDLINE=94008528; PubMed=8404528;
RA Whittaker C.A., Desimone D.W.;
RT "Integrin alpha subunit mRNAs are differentially expressed in early Xenopus embryos";
RT Xenopus embryos.;
RL Development 117:1239-1249 (1993).
DR EMBL; L43058; AAA69770.1; -;
DR EMBL; L10186; AAA16246.1; -;
DR PIR; I51524; I51524.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007140; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR Pfam; PF01839; FG-GAP; 3.
DR Pfam; PF00357; Integrin_A; 1.
DR PRINTS; PR01185; INTEGRIN.
DR SMART; SM00191; Int_alpha; 4.
DR PROSITE; PS00242; INTEGRIN_ALPHA; 1.
FT NON TER 1
SQ SEQUENCE 780 AA; 87017 MW; 9518B18C2B6BF637 CRC64;

Query Match 0.8%; Score 9; DB 13; Length 780;
Best Local Similarity 100.0%; Pred.No.12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGP 503
DB 108 TDVLLVGP 116

RESULT 16

ID Q9TS65 PRELIMINARY; PRT; 54 AA.

AC Q9TS65;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE LEUKOINTEGRIN alpha D beta 2 (Fragments).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE.
RX MEDLINE=95325609; PubMed=7541420;
RA Danilenko D.M., Rossitto P.V., Van der Vieren M., Le Tronq H.,

```

RA McDonough S.P., Affolter V.K., Moore P.F.;
RT "A novel canine leukointegrin, alpha d beta 2, is expressed by
RT specific macrophage subpopulations in tissue and a minor CD8+
RT lymphocyte subpopulation in peripheral blood.";
PL J. Immunol. 155:35-44(1995).
FT NON_TER 1
FT NON_CONS 23
FT NON_TER 24
FT NON_TER 54
SQ SEQUENCE 54 AA; 5624 MW; C2FC82C2FDEC3E6C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 54;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 24 LVVGAPLE 31

RESULT 17
Q88NW3 ID Q88NW3 PRELIMINARY; PRT; 58 AA.
AC Q88NW3;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Conserved hypothetical protein.
OS Pseudomonas
GN Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=2243060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L., Beanan M., DeBoy R.T., Daugherty S., Kolonay J.,
RA Madupu R., Nelson W., White O., Peterson J., Khouri H., Hance I.,
RA Chris Lee P., Holtzaple E., Scanlan D., Tran K., Moazzez A.,
RA Uterback T., Rizzo M., Lee K., Kosack D., Moestl D., Wedler H.,
RA Lauber J., Stjepandic D., Hohnselt J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J., Timmis K.N., Duesterhoeft A., Tuemmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440.";
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016778; AAN66716.1; -.
DR TIGR; PP1091; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 58 AA; 6734 MW; CD9350CEF793D269 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 884 EERLQKQ 891
Db 43 EERLQKQ 50

RESULT 18
Q8CEU3 ID Q8CEU3 PRELIMINARY; PRT; 76 AA.
AC Q8CEU3;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE CD24a antigen.
GN CD24A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=Head;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK014279; BAC25429.1; -.
DR PIR; PT0721; PT0721.
DR MGD; MGI:88323; Cd24a.
SQ SEQUENCE 76 AA; 7787 MW; 6853F12240F91AEB CRC64;

Query Match 0.7%; Score 8; DB 11; Length 76;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 19
Q9N241 ID Q9N241 PRELIMINARY; PRT; 78 AA.
AC Q9N241;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Saimiri boliviensis (Bolivian squirrel monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Saimiri.
OX NCBI_TaxID=27679;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125670; AAF64385.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin...; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 78
SQ SEQUENCE 78 AA; 7607 MW; D4E684AD6F6375B CRC64;

Query Match 0.7%; Score 8; DB 6; Length 78;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 64 LLLALLVL 71

RESULT 20
Q9N242 ID Q9N242 PRELIMINARY; PRT; 87 AA.
AC Q9N242;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)

```

```

DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Ateles sp. (Spider monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Ateleinae; Ateles.
OX NCBI_TaxID=9311;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38 (2000).
DR EMBL; AF125669; AAF64384.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT NON TER
SQ SEQUENCE 87 AA; 8388 MW; 64574359DAC555D4 CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 87;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 73 LLLALLVL 80
|||||
|||||

RESULT 21
Q9587 PRELIMINARY; PRT; 88 AA.
AC Q9587;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DE Dopamine D4 receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96362071; PubMed=8746407;
RA Cichon S., Nothen M.M., Catalano M., Di Bella D., Maier W.,
RA Lichtermann D., Minges J., Albus M., Bormann M., Franzek E., et al;
RT "Identification of two novel polymorphisms and a rare deletion variant
RT in the human dopamine D4 receptor gene.";
RL Psychiatr. Genet. 5:97-103 (1995).
DR EMBL; S82918; AAB46803.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT NON TER
SQ SEQUENCE 88 AA; 8482 MW; 314093AB4678CEDF CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 4; Length 88;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161

```

```

Db 74 LLLALLVL 81
|||||
|||||

RESULT 22
Q9N244 PRELIMINARY; PRT; 91 AA.
AC Q9N244;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Nasalis larvatus (Proboscis monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Colobinae;
OC Nasalis.
OX NCBI_TaxID=43780;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38 (2000).
DR EMBL; AF125667; AAF64382.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT NON TER
SQ SEQUENCE 91 AA; 8827 MW; 67B53D0764BD439C CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 77 LLLALLVL 84
|||||
|||||

RESULT 23
Q9N245 PRELIMINARY; PRT; 91 AA.
AC Q9N245;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38 (2000).
DR EMBL; AF125666; AAF64381.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.

```

```

DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 91 AA; 8728 MW; 7CFF8CAC7E0CF39C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 77 LLLALLVL 84

RESULT 24
ID O97517 PRELIMINARY; PRT; 91 AA.
AC O97517;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
DR DRD4.
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010297; AAC67225.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 91 AA; 8851 MW; 9DF2F56F1D7CF828 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 77 LLLALLVL 84

RESULT 25
O97515 PRELIMINARY; PRT; 91 AA.
AC O97515;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
DR DRD4.
OS Pan paniscus (Pygmy chimpanzee) (Bonobo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9597;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010295; AAC67223.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 91 AA; 8763 MW; 76BDBF6BE386E299 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 77 LLLALLVL 84

RESULT 26
O97521 PRELIMINARY; PRT; 91 AA.
AC O97521;
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
DR DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010301; AAC67229.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; P:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
SQ SEQUENCE 91 AA; 8934 MW; 93CD0F1CFC8EBE2C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 91;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 77 LLLALLVL 84

RESULT 27
O9N243 PRELIMINARY; PRT; 91 AA.
ID O9N243

```

```

AC Q9N243;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Presbytis cristata (Silvered langur).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Colobinae;
OC Presbytis.
OX NCBI_TaxID=36232;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF125668; AA64383.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT SEQUENCE 91 AA; 8827 MW; 67E53D0764BD439C CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 77 LLLALLVL 84
|||||

RESULT 28
Q9N246 PRELIMINARY; PRT; 91 AA.
AC Q9N246;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Papio.
OX NCBI_TaxID=9557;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF125665; AA64380.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT SEQUENCE 91 AA; 8770 MW; 608F90C00E10839C CRC64;

```

```

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 91;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 77 LLLALLVL 84
|||||

RESULT 29
Q9N247 PRELIMINARY; PRT; 93 AA.
AC Q9N247;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates syndactylus (Siamang) (Symphalangus syndactylus).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9590;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF125664; AA64379.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER
FT SEQUENCE 93 AA; 8904 MW; 5D07B793F2B624E4 CRC64;

Query Match
Best Local Similarity 0.7%; Score 8; DB 6; Length 93;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 79 LLLALLVL 86
|||||

RESULT 30
Q99586 PRELIMINARY; PRT; 95 AA.
AC Q99586;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96382071; PubMed=8746407;
RA Cichon S., Nothen M.M., Catalano M., Di Bella D., Maier W.,
RA Lichtermann D., Minges J., Albus J., Bormann M., Franzek E., et al;
RT "Identification of two novel polymorphisms and a rare deletion variant
RT in the human dopamine D4 receptor gene.";
RL Psychiatr. Genet. 5:97-103(1995).
DR EMBL: S82917; AAB46802.1; -.

```


DR GO: GO:0004952; F:dopamine receptor activity; NAS.
DR GO: GO:0007212; P:dopamine receptor signaling pathway; NAS.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9148 MW; 7163AF13B6156686 CRC64;

Query Match 0.7%; Score 8; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 31
O97514 PRELIMINARY; PRT; 95 AA.
AC O97514; 95 AA.
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010294; AAC67222.1; .
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9119 MW; 02BD95057095863C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 32
O9N249 PRELIMINARY; PRT; 95 AA.
AC O9N249; 95 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates klossii (Kloss's gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF010294; AAC67222.1; .
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9119 MW; 02BD95057095863C CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 33
O9N248 PRELIMINARY; PRT; 95 AA.
AC O9N248; 95 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125663; AAF64378.1; .
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 34
O9N248 PRELIMINARY; PRT; 95 AA.
AC O9N248; 95 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125663; AAF64378.1; .
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 34

OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125662; AAF64377.1; .
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 33
O9N248 PRELIMINARY; PRT; 95 AA.
AC O9N248; 95 AA.
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Hylobates lar (Common gibbon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hylobatidae; Hylobates.
OX NCBI_TaxID=9580;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinard A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL; AF125663; AAF64378.1; .
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm 1; 1.
DR PRINTS: PR00237; GPCR_Rhodopsn.
DR PROSITE: PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 95 95
SQ SEQUENCE 95 AA; 9090 MW; 47BC27C372E1F12F CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 34

```
O97516
ID O97516 PRELIMINARY; PRT; 95 AA.
AC O97516;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Gorilla gorilla (Gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinaid A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010296; AAC67224.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER
FT NON_TER
SQ SEQUENCE 95 AA; 9121 MW; ECF6C50B3A809D33 CRC64;

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred.No.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 35
O97520
ID O97520 PRELIMINARY; PRT; 95 AA.
AC O97520;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinaid A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010296; AAC67228.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0004872; F:receptor activity; IEA.
DR GO: GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P:G-protein coupled receptor protein signalin. .; IEA.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor.
FT NON_TER
FT NON_TER
SQ SEQUENCE 95 AA; 9295 MW; 1C5CD4013C3B365E CRC64;
```

```
O97518
ID O97518 PRELIMINARY; PRT; 95 AA.
AC O97518;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinaid A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010298; AAC67226.1; -.

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred.No.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 37
O97518
ID O97518 PRELIMINARY; PRT; 99 AA.
AC O97518;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Dopamine D4 receptor (Fragment).
GN DRD4.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20215006; PubMed=10750051;
RA Seaman M.I., Chang F.M., Deinaid A.S., Quinones A.T., Kidd K.K.;
RT "Evolution of exon 1 of the dopamine D4 receptor (DRD4) gene in
RT primates.";
RL J. Exp. Zool. 288:32-38(2000).
DR EMBL: AF010298; AAC67226.1; -.

Query Match 0.7%; Score 8; DB 6; Length 95;
Best Local Similarity 100.0%; Pred.No.21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88
```

```
DR GO: GO:0016021; C: integral to membrane; IEA.
DR GO: GO:004872; F: receptor activity; IEA.
DR GO: GO:001584; F: rhodopsin-like receptor activity; IEA.
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro: IPR000216; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
FT NON TER          99      99
SQ SEQUENCE 99 AA; 9335 MW; D4CED9750184BC47 CRC64;

Query Match          0.7%; Score 8; DB 6; Length 99;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 85 LLLALLVL 92

RESULT 38
Q8E9P1 PRELIMINARY; PRT; 104 AA.
AC Q8E9P1;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cell division protein FtsL.
GN FtsL OR S04226.
OS Shewanella oneidensis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC Alteromonadaceae; Shewanella.
OX NCBI_TaxID=70863;
RN [1]
SQ SEQUENCE FROM N.A.
STRAIN=MR-1;
MEDLINE=22297686; PubMed=12368813;
RA Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA Read T.D., Eisen J.A., Seshadri R., Ward N., Methe B., Clayton R.A.,
RA Meyer T., Taapin A., Scott J., Beanan M., Brinkac L., Daugherty S.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Haft D.H., Kolonay J.F.,
RA Madupu R., Peterson J.D., Umayam L.A., White O., Wolf A.M.,
RA Vamathevan J., Weidman J., Impraim M., Lee K., Berry K., Lee C.,
RA Mueller J., Khouri H., Gill J., Utterback T.R., McDonald L.A.,
RA Feldblyum T.V., Smith H.O., Venter J.C., Neilson K.H., Fraser C.M.;
RA "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT Shewanella oneidensis."
RL Nat. Biotechnol. 20:1118-1123 (2002).
DR EMBL; AE015855; AAN57198.1; -.
DR TIGR; SO4226; -.
DR InterPro; IPR007082; FtsL.
DR Pfam; PF04999; FtsL; 1.
DR Complete proteome.
KW Complete proteome.
SQ SEQUENCE 104 AA; 12187 MW; 408B210954032920 CRC64;

Query Match          0.7%; Score 8; DB 16; Length 104;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 23 LLLALLVL 30

RESULT 39
Q9SDS4 PRELIMINARY; PRT; 106 AA.
AC Q9SDS4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Non-specific lipid transfer protein.

GN LPTI.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
SQ SEQUENCE FROM N.A.
STRAIN=cv._Hanbyul; TISSUE=Leaf;
RA Hwang B.K., Jung H.W.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208832; AAF23458.1; -.
DR HSSP; P23096; IRZL.
DR GO: GO:0008289; F: lipid binding; IEA.
DR GO: GO:0006869; P: lipid transport; IEA.
DR InterPro; IPR003612; AAI.
DR InterPro; IPR000528; Plant LTP.
DR Pfam; PF00234; tryptophan amyl; 1.
DR PRINTS; PR00382; LIPIDTRANSF.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 106 AA; 11291 MW; 5E9C43238B9381C9 CRC64;

Query Match          0.7%; Score 8; DB 10; Length 106;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 197 GQIQGVGV 204
Db 20 GQIQGVGV 27

RESULT 40
Q8R1E9 PRELIMINARY; PRT; 112 AA.
AC Q8R1E9;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN C2A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SQ SEQUENCE FROM N.A.
TISSUE=Salivary gland;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024677; AAH24677.1; -.
DR MGD; MGI:2655426; C2a.
DR InterPro; IPR006038; Uteroglobin_supf.
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 112 AA; 12371 MW; 69CE957825183F72 CRC64;

Query Match          0.7%; Score 8; DB 11; Length 112;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 LLLALLVL 1160
Db 5 LLLALLVL 12

RESULT 41
Q8JZX1 PRELIMINARY; PRT; 112 AA.
AC Q8JZX1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein (Salivary androgen-binding protein gamma
DE subunit).
```

C2D OR ABPG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1) _____
 RP SEQUENCE FROM N.A.
 RC TISSUE=Salivary gland;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN (2) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeJ;
 RX STRAIN=C3H/HeJ;
 EX MEDLINE=22680631; PubMed=12795612;
 RA Karn R.C., Laukaitis C.M.;
 RT "Characterization of two forms of mouse salivary androgen-binding
 protein (ABP): implications for evolutionary relationships and ligand-
 binding function.";
 RT Biochemistry 42:7162-7170(2003).
 RN (3) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeJ;
 RA Laukaitis C.M., Dlouhy S.R., Karn R.C.;
 RT "The mouse salivary androgen-binding protein (ABP) gene cluster on
 Chromosome 7: Characterization and evolutionary relationships.";
 RL Mamm. Genome 0:0-0(2003).
 DR EMBL; BC035222; AAH35222.1; -.
 DR EMBL; AY325898; AAP94990.1; -.
 DR MGD; MGI:2655745; C2d.
 KW Hypothetical protein.
 SQ SEQUENCE 112 AA; 12794 MW; F83285868F4BD53D CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
 DB 5 LLLALLV 12

RESULT 42

Q7TNV5 PRELIMINARY; PRT; 112 AA.
 AC Q7TNV5;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DE Salivary androgen-binding protein beta subunit.
 GN ABPG.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN (1) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeJ;
 RX MEDLINE=22680631; PubMed=12795612;
 RA Karn R.C., Laukaitis C.M.;
 RT "Characterization of two forms of mouse salivary androgen-binding
 protein (ABP): implications for evolutionary relationships and ligand-
 binding function.";
 RT Biochemistry 42:7162-7170(2003).
 RN (2) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=C3H/HeJ;
 RA Laukaitis C.M., Dlouhy S.R., Karn R.C.;
 RT "The mouse salivary androgen-binding protein (ABP) gene cluster on
 Chromosome 7: Characterization and evolutionary relationships.";
 RL Mamm. Genome 0:0-0(2003).
 DR EMBL; AY325897; AAP94989.1; -.
 SQ SEQUENCE 112 AA; 12383 MW; EB5F2C68BFB8D17C CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
 DB 5 LLLALLV 12

RESULT 43

Q7TNN0 PRELIMINARY; PRT; 112 AA.
 AC Q7TNN0;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DE Salivary androgen-binding protein beta subunit.
 GN ABPG.
 OS Mus musculus musculus (eastern European house mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=39442;
 RN (1) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=DBA/2J;
 RX MEDLINE=22680631; PubMed=12795612;
 RA Karn R.C., Laukaitis C.M.;
 RT "Characterization of two forms of mouse salivary androgen-binding
 protein (ABP): implications for evolutionary relationships and ligand-
 binding function.";
 RT Biochemistry 42:7162-7170(2003).
 DR EMBL; AY293278; AAP44465.1; -.
 SQ SEQUENCE 112 AA; 12411 MW; 818431A8CEFC029 CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
 DB 5 LLLALLV 12

RESULT 44

Q7TNW7 PRELIMINARY; PRT; 112 AA.
 AC Q7TNW7;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DE Salivary androgen-binding protein gamma subunit.
 GN ABPG.
 OS Mus musculus musculus (eastern European house mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=39442;
 RN (1) _____
 RP SEQUENCE FROM N.A.
 RC STRAIN=DBA/2J;
 RX MEDLINE=22680631; PubMed=12795612;
 RA Karn R.C., Laukaitis C.M.;
 RT "Characterization of two forms of mouse salivary androgen-binding
 protein (ABP): implications for evolutionary relationships and ligand-
 binding function.";
 RT Biochemistry 42:7162-7170(2003).
 DR EMBL; AY293281; AAP44468.1; -.
 SQ SEQUENCE 112 AA; 12546 MW; 9F955317D9116883 CRC64;

Query Match 0.7%; Score 8; DB 11; Length 112;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160

```
Db          |||||
           5 LLLLLLV 12

RESULT 45
Q9ZPI9
ID Q9ZPI9 PRELIMINARY; PRT; 114 AA.
AC Q9ZPI9;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Lipid transfer protein LTP1.
GN LTP1.
OS Capsicum annuum (Bell pepper).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Capsicum.
OX NCBI_TaxID=4072;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. VK-1;
RA Park C.J., Shin R., Paek K.H.;
RT "Characterization of a lipid transfer protein (CaLTP1) in pepper.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF118131; AAD18029.1; -.
DR HSSP; P23096; LRZL.
DR GO; GO:0008289; F:lipid binding; IEA.
DR GO; GO:0006869; P:lipid transport; IEA.
DR InterPro; IPR003612; AAI.
DR DR Pfam; PF00234; try_alpha_amyl; 1.
DR PRINTS; PR00382; LIPIDTRANSFES.
DR SMART; SM00499; AAI; 1.
SQ SEQUENCE 114 AA; 12114 MW; 9315BA2125A83012 CRC64;

Query Match 0.7%; Score 8; DB 10; Length 114;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 197 GQIQGVV 204
    |||||
Db 28 GQIQGVV 35

RESULT 46
Q9RSB5
ID Q9RSB5 PRELIMINARY; PRT; 127 AA.
AC Q9RSB5;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein DR2210.
GN DR2210.
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Hart D.H., Gunn M.L., Nelson W.C., Richardson D.L.,
RA Morfat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vanathevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE002054; AAF11763.1; -.
DR PIR; B75301; B75301.
```

```

RT 60,770 full-length cDNAs."
RL Nature 420:563-573(2002).
DR EMBL; AK083120; BAC38769.1; -.
DR PIR; PT0714; PT0714.
DR MGI; MGI:107742; Stra6.
SQ SEQUENCE 137 AA; 14573 MW; 27A54C974EF9670C CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 11; Length 137;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
DB 64 LLLALLLV 71

RESULT 49
Q61692 PRELIMINARY; PRT; 141 AA.
ID Q61692
AC Q61692;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative heat stable antigen.
GN HSA-C.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-CBA X C57BL/6; TISSUE=Spleen;
RX MEDLINE=91209380; PubMed=2019286;
RA Wenger R.H., Ayane M., Bose R., Koehler G., Nielsen P.J.;
RT "The genes for a mouse hematopoietic differentiation marker called the
RT heat-stable antigen."
RL Eur. J. Immunol. 21:1039-1046(1991).
DR EMBL; X56486; CNA3843.1; -.
DR PIR; S15785; S15785.
SQ SEQUENCE 141 AA; 15515 MW; E4BFB428ADC03C69 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 11; Length 141;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 12 GLLLLALL 19

RESULT 50
Q71LM3 PRELIMINARY; PRT; 144 AA.
ID Q71LM3
AC Q71LM3;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Bradykinin-potentiating/C-type natriuretic peptide isoform 2
DE (Fragment)
OS Bothrops jararacussu (Jararacussu).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8726;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Lameu C., Radis-Baptista G., Yamane T.,
RA Camargo A.C.M.;
RT "Cloning and sequence analysis of a Bothrops jararacussu Bpps
RT precursor."
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY310915; AAP83422.1; -.

FT NON_TER 144 144
SQ SEQUENCE 144 AA; 15982 MW; B23FED5230A84910 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 13; Length 144;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 10 GLLLLALL 17

RESULT 51
Q8QG90 PRELIMINARY; PRT; 146 AA.
ID Q8QG90
AC Q8QG90;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Bradykinin-potentiating protein.
OS Bothrops insularis (Island jararaca) (Queimada jararaca)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidodactylus; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8723;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Junqueira-de-Azevedo I.L.M., Ho P.L.;
RT "A survey of gene expression and diversity in the venom glands of the
RT pitviper Bothrops insularis through the generation of Expressed
RT Sequence Tags (ESTs)."
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490532; AAM09691.1; -.
SQ SEQUENCE 146 AA; 16088 MW; F58244C4F24B1799 CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 13; Length 146;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
DB 10 GLLLLALL 17

RESULT 52
Q7XK40 PRELIMINARY; PRT; 161 AA.
ID Q7XK40
AC Q7XK40;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSJNB0044K18.5 protein.
GN OSJNB0044K18.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Fu G., Wang S.Y., Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F.,
RA Jia J., Yin H.F., Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Ding C.W.,
RA Shao Y., Sun Y., Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Feng C.W.,
RA Sheng H.H., Gu J.L., Chen S.T., Ni L., Zhu F.H., Han B., Feng Q.,
RA Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X., Liu Y.L., Mu J., Yu Z.,
RA Chen L., Fan D.L., Weng Q.J., Zhang L., Lu Y.Q., Yu S.L., Liu X.H.,
RA Lu T.T., Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
RA Qian Y.M., Ying K., Zhou B., Chen Z.H., Hao P., Zhang L., Wu M.,
RA Zhang R.Q., Guan J.P., Hong G.F.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL606595; CAE05863.1; -.
SQ SEQUENCE 161 AA; 16300 MW; 0E8B2DF9C58ADC61 CRC64;

```

```

Query Match          0.7%; Score 8; DB 10; Length 161;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1156 LALLVLAL 1163
Db 136 LALLVLAL 143

RESULT 53
Q98MM8 PRELIMINARY; PRT; 174 AA.
AC Q98MM8;
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-NAR-2002 (TRENBLrel. 20, Last annotation update)
DE Hypothetical protein mlr0512.
GN MLR0512.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338 (2000).
DR EMBL; AP002995; BAB48085.1; -;
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 174 AA; 20270 MW; E93E12118C11022F CRC64;

Query Match          0.7%; Score 8; DB 16; Length 174;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1090 LRSALK 1097
Db 38 LRSALK 45

RESULT 54
P79786 PRELIMINARY; PRT; 181 AA.
AC P79786;
DT 01-MAY-1997 (TRENBLrel. 03, Created)
DT 01-MAY-1997 (TRENBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Thrombospondin-1 (Fragment).
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97209337; PubMed=9056637;
RA Tucker R.P., Haglos C., Chiquet-Ehrismann R., Lawler J.;
RT "In situ localization of thrombospondin-1 and thrombospondin-3
RT transcripts in the avian embryo.";
RL Dev. Dyn. 208:328-337 (1997).
DR EMBL; U76994; AAB19208.1; -;
DR GO; GO:0005509; F:calcium ion binding; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR InterPro; IPR003367; tsp_3.
DR Pfam; PF02412; tsp_3; 8.

Query Match          0.7%; Score 8; DB 13; Length 181;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17

RESULT 56
Q90Y11 PRELIMINARY; PRT; 181 AA.
AC Q90Y11;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Bradykinin potentiating peptide and C-type natriuretic peptide isoform
DE 2.
OS Crotales durissus terrificus (South American rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8732;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Radis-Baptista G., Barbosa S.R., Yamane T.,
RA Camargo A.C.M.;
RT "Crotales durissus terrificus bradykinin potentiating peptide
RT precursor";
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF308593; AAL09426.1; -;
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP.1.
DR PROSITE; PS00263; NATRIURITIC_PEPTIDE; 1.
SQ SEQUENCE 181 AA; 18560 MW; 7B5ADC5B9372D07F CRC64;

Query Match          0.7%; Score 8; DB 13; Length 181;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17

RESULT 56
Q90Y11 PRELIMINARY; PRT; 181 AA.
AC Q90Y11;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE Bradykinin potentiating peptide and C-type natriuretic peptide isoform
DE 2.
OS Crotales durissus terrificus (South American rattlesnake).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Crotalus.
OX NCBI_TaxID=8732;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Venom gland;
RA Hayashi M.A.F., Radis-Baptista G., Barbosa S.R., Yamane T.,
RA Camargo A.C.M.;

```

RT "Crotalus durissus terrificus bradykinin-potentiating peptide and C-
 RT type natriuretic peptide precursor isoform2.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF308594; AAL09427.1; -
 DR GO; GO:0005576; C:extracellular; IEA.
 DR GO; GO:0005179; F:hormone activity; IEA.
 DR InterPro; IPR000863; Natri_peptide.

DR Pfam; PF00212; ANP; 1.
 DR PRINTS; PRO0710; NATPEPTIDES.
 DR SMART; SM00183; NAT_PEP; 1.
 DR PROSITE; PS00263; Natriuretic_Peptide; 1.

SQ SEQUENCE 181 AA; 18507 MW; 982B95D38AASF27 CRC64;
 Query Match 0.7%; Score 8; DB 13; Length 181;
 Best Local Similarity 100.0%; Pred.No. 37;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159

DB 10 GLLLLALL 17

RESULT 57

Q8U4N4

ID Q8U4N4

AC Q8U4N4

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)

DE Hypothetical protein PF0047.

GN PF0047.

OS Pyrococcus furiosus

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI_TaxID=2261;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=V61 / DSM 3638 / ATCC 43587 / JCM 8422;

RA Weiss R.B., Dunn D.M., Robb P.T., Brown J.R.;

RT "The complete sequence of the Pyrococcus furiosus genome.";

RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF010131; AAL080171.1; -

DR Hypothetical protein; Complete proteome.

SW SEQUENCE 222 AA; 24556 MW; 721D79ED547A0F13 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 222;

Best Local Similarity 100.0%; Pred.No. 44;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 RGKVVYVE 519

DB 106 RGKVVYVE 113

RESULT 58

Q7V5L4

ID Q7V5L4

AC Q7V5L4

DT 01-OCT-2003 (TrEMBLrel. 25, Created)

DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)

DE Conserved hypothetical membrane protein.

GN F3H9.20 OR PMT1539.

OS Prochlorococcus marinus (strain MIT 9313).

OC Bacteria; Cyanobacteria; Prochlorophytes; Prochlorococcaceae;

OC Prochlorococcus.

OX NCBI_TaxID=74547;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=22825698; PubMed=12917642;

RA Rocap G., Larimer F.W., Lamerdin J., Malfatti S., Chain P.,

RA Ahlgren N.A., Arellano A., Coleman M., Hauser L., Hess W.R.,

RA Johnson Z.I., Land M., Lindell D., Post A.F., Regala W., Shah M.,

RA Shaw S.L., Steglich C., Sullivan M.B., Ting C.S., Tolonen A.,
 RA Webb E.A., Zinser E.R., Chisholm S.W.;

RT "Genome divergence in two prochlorococcus ecotypes reflects oceanic
 RT niche differentiation.";

RL Nature 424:1042-1047(2003).

DR EMBL; BX572099; CAE21714.1; -

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 224 AA; 24172 MW; E1CD9F149C279A29 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 224;

Best Local Similarity 100.0%; Pred.No. 44;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1158 LLVTLARK 1165

DB 190 LLVTLARK 197

RESULT 59

O57835

ID O57835

AC O57835

DT 01-AUG-1998 (TrEMBLrel. 07, Created)

DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)

DE Hypothetical protein PH0087.

GN PH0087.

OS Pyrococcus horikoshii.

OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;

OC Pyrococcus.

OX NCBI_TaxID=53953;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=OT3;

RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,

RA Yamamoto S., Sekine M., Baba S.-I., Koguchi H., Hosoyama A., Nagai Y.,

RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,

RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A.,

RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb P.T., Horikoshi K.,

RA Masuchi Y., Shizuya H., Kikuchi H.

RT "Complete sequence and gene organization of the genome of a hyper-

RT thermophilic archaebacterium, Pyrococcus horikoshii OT3.";

RL DNA Res. 5:55-76(1998).

DR EMBL; AP000001; BAA29156.1; -

DR PIR; E71228; E71228.

KW Hypothetical protein; Complete proteome.

SQ SEQUENCE 224 AA; 25134 MW; 9CC56212064558E0 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 224;

Best Local Similarity 100.0%; Pred.No. 44;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 512 RGKVVYVE 519

DB 105 RGKVVYVE 112

RESULT 60

P74604

ID P74604

AC P74604

DT 01-FEB-1997 (TrEMBLrel. 02, Created)

DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)

DE Hypothetical protein slr1571.

GN SLR1571.

OS Synechocystis sp. (strain PCC 6803).

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX NCBI_TaxID=1148;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97061201; PubMed=8905231;

RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hirosewa M., Sugura M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
 RA Shingo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
 RA Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT *Synechocystis* sp. strain PC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 DR EMBL; D90916; BAA18712.1; -;
 DR PIR; S76800; S76800.
 DR InterPro; IPR002838; DUF124.
 DR Pfam; PF01987; DUF124; 1.
 DR ProDom; PD013634; DUF124; 1.
 DR TIGRFam; TIGR00266; TIGR00266; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 226 AA; 25003 MW; 29C5C137C128DAC0 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 226;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1073 RLVPNCEI 1080
 DB 183 RLVPNCEI 190

RESULT 61
 ID Q882P2 PRELIMINARY; PRT; 227 AA.
 AC Q882P2;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Conserved hypothetical protein.
 GN PSP02584.
 OS *Pseudomonas syringae* (pv. tomato).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=323;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=DC3000;
 RA Buell R., Joardar V., Khouri H., Fedorova N., Tran B., Russell D.,
 RA Berry K., Utterback T., Van Aken S., Feldblyum T., Gwinn M.,
 RA Dodson R., DeBoy R., Durkin A., Kolonay J., Madupu R., Daugherty S.,
 RA Brinkac L., Beanan M., Haft D., Selengut J., Nelson W., Davidson T.,
 RA White O., Fraser C., Collier A.;
 RT "Complete sequence of *Pseudomonas syringae*.";
 RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE016865; AAC056088.1; -;
 DR TIGR; PSP02584; -;
 DR InterPro; IPR002489; DUF14.
 DR Pfam; PF01493; GXGXG; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 227 AA; 24036 MW; AB3458F347BBF145 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 227;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 547 ASVRDLNQ 554
 DB 9 ASVRDLNQ 16

RESULT 62
 ID Q9AJX4 PRELIMINARY; PRT; 231 AA.
 AC Q9AJX4;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)

DE Putative oxidoreductase.
 GN SC01803 OR SCI33.02.
 OS Streptomyces coelicolor.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Harris D.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kieser H.M., Denapaita D., Eichner A., Cullum J.,
 RA Kinashi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
 RL Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2) / M145;
 RX MEDLINE=21996410; PubMed=12000953;
 RA Bentley S.D., Chater K.F., Cardeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Gobie A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kieser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 RT coelicolor A3(2).";
 RL Nature 417:141-147(2002).
 CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES
 (SDR) FAMILY.
 CC EMBL; AL939110; CAC28543.1; -;
 DR HSP; O64105; 10AA.
 DR GO; GO:0016491; F:oxidoreductase activity; IEA.
 DR GO; GO:0008152; P:metabolism; IEA.
 DR InterPro; IPR002198; ADH_short.
 DR Pfam; PF00106; adh_short; 1.
 DR PRINTS; PR00080; SDRFAMILY.
 DR PROSITE; PS00061; ADH_SHORT; 1.
 KW Oxidoreductase; Complete proteome.
 SQ SEQUENCE 231 AA; 23789 MW; 6AEDEB40306B94086 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 231;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 790 DLVLDARS 797
 DB 27 DLVLDARS 34

RESULT 63
 ID Q98GF8 PRELIMINARY; PRT; 235 AA.
 AC Q98GF8;
 DT 01-OCT-2001 (TrEMBLrel. 18, Created)
 DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE ATP-binding component of phosphonate transport, HisP-like
 DE nucleotide binding protein, PhnL.
 GN MLR3349.

OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 OC Phyllobacteriaceae; Mesorhizobium.
 ON NCBI_TaxID=381;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAFF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
 RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti";
 RL DNA Res. 7:331-338(2000).
 DR EMBL; AP003001; BAB50258.1; --
 DR GO; GO:0015020; C:membrane; IEA.
 DR GO; GO:0005524; F:ATP binding; IEA.
 DR GO; GO:0004009; F:ATP-binding cassette (ABC) transporter acti. . .; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR003439; ABC transporter.
 DR Pfam; PF00005; ABC tran_1.
 DR ProDom; PD000006; ABC transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR PROSITE; PS00211; ABC_TRANSPORTER_1; 1.
 DR PROSITE; PS00893; ABC_TRANSPORTER_2; 1.
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 235 AA; 25098 MW; 6FA8F03F827361 CRC64;
 Query Match 0.7%; Score 8; DB 16; Length 235;
 Best Local Similarity 100.0%; Pred.No. 46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 616 EDGLIDLA 623
 DB 73 EDGLIDLA 80
 RESULT 64
 Q20220 PRELIMINARY; PRT; 245 AA.
 AC Q20220;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein F40F4.7.
 GN F40F4.7.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
 OC Rhabditidae; Felodierinae; Caenorhabditis.
 ON NCBI_TaxID=6239;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RX MEDLINE=99069613; PubMed=9851916;
 RA Waterston R.;
 RT "Genome sequence of the nematode C. elegans: a platform for
 RT investigating biology. The C. elegans Sequencing Consortium.";
 RL Science 282:2012-2018(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Wilson R.;
 RT "The sequence of C. elegans cosmid F40F4.";
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Bristol N2;
 RA Waterston R.;
 RT Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; U40420; AAK84526.3; --
 DR WormPep; F40F4.7; CE28316.
 DR GO; GO:0008080; P:N-acetyltransferase activity; IEA.
 DR InterPro; IPR000182; GCSacetyl_trans.
 DR Pfam; PF00583; Acetyltransf; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 245 AA; 27893 MW; A53289FBB2F91E75 CRC64;
 Query Match 0.7%; Score 8; DB 5; Length 245;
 Best Local Similarity 100.0%; Pred.No. 48;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 557 YNDVVVGA 564
 DB 141 YNDVVVGA 148
 RESULT 65
 Q8Y199 PRELIMINARY; PRT; 257 AA.
 AC Q8Y199;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein RSC0620.
 GN RSC0620 OR RS01521.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
 OC Burkholderiaceae; Ralstonia.
 ON NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GMT1000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Ariat M., Billault A., Brottier P., Camus J.C., Catolico L.,
 RA Chardier M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schiex T.,
 RA Sigulier P., Thebault P., Whalen M., Wincker P., Levy M.,
 RA Weissenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
 RL Nature 415:457-502(2002).
 DR EMBL; AL646060; CAD1450.1; --
 DR GO; GO:0006118; P:electron transport; IEA.
 DR InterPro; IPR000572; Oxidored_molyb.
 DR Pfam; PF00174; oxidored_molyb; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 257 AA; 29260 MW; DAE569BE1572113D CRC64;
 Query Match 0.7%; Score 8; DB 16; Length 257;
 Best Local Similarity 100.0%; Pred.No. 50;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1149 TLGGLLLL 1156
 DB 37 TLGGLLLL 44
 RESULT 66
 Q8QG91 PRELIMINARY; PRT; 265 AA.
 AC Q8QG91;
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Bradykinin-potentiating/c-type natriuretic protein.
 OS Bothrops insularis (Island jararaca) (Queimada jararaca).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
 OC Viperidae; Crotalinae; Bothrops.
 ON NCBI_TaxID=8723;
 RN [1]
 RP SEQUENCE FROM N.A.

RC TISSUE=Venom gland;
RA Junqueira-de-Azevedo I.L.M., Ho P.L.;
RT "A survey of gene expression and diversity in the venom glands of the
RT pitviper Bothrops insularis through the generation of Expressed
RT Sequence Tags (ESTs)."; the EMBL/GenBank/DBJ databases.
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF490531; AAM09690.1; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
SQ SEQUENCE 265 AA; 27763 MW; 0EAE1408B42358BE CRC64;

Query Match 0.7%; Score 8; DB 13; Length 265;
Best Local Similarity 100.0%; Pred.No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17
|||||
|

RESULT 67
Q9PW56 PRELIMINARY; PRT; 265 AA.
AC Q9PW56;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Bradykinin-potentiation/C-type natriuretic peptide.
OS Bothrops jararaca (Jararaca).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Bothrops.
OX NCBI_TaxID=8724;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Hayashi M.A.F., Murbach A.F., Camargo A.C.M.;
RT "The precursor of C-type natriuretic peptide of snake brain contains
RT angiotensin converting enzyme inhibitors, specific for the C-catalytic
RT site".
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF171670; AAD51326.2; -
DR GO; GO:0005576; C:extracellular; IEA.
DR GO; GO:0005179; F:hormone activity; IEA.
DR InterPro; IPR000663; Natr_peptide.
DR Pfam; PF00212; ANP; 1.
DR PRINTS; PR00710; NATPEPTIDES.
DR SMART; SM00183; NAT_PEP; 1.
DR PROSITE; PS00263; NATRIURETIC_PEPTIDE; 1.
SQ SEQUENCE 265 AA; 27763 MW; 8E99AEC976CCD439 CRC64;

Query Match 0.7%; Score 8; DB 13; Length 265;
Best Local Similarity 100.0%; Pred.No. 51;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 10 GLLLLALL 17
|||||
|

RESULT 68
Q861M5 PRELIMINARY; PRT; 272 AA.
AC Q861M5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MHC class I antigen (Fragment).

OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=107;
RA Chung C., Leib S.R., Fraser D.G., Ellis S.A., McGuire T.C.;
RT "Sequence-based typing of horse classical MHC class I genes.";
RT Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY176095; AAO18112.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG-cl.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; Igc1; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT NON_TER 272
SQ SEQUENCE 272 AA; 30576 MW; 8EC9E0C183D319FD CRC64;

Query Match 0.7%; Score 8; DB 7; Length 272;
Best Local Similarity 100.0%; Pred.No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VLLGAVGA 387
Db 232 VLLGAVGA 239
|||||
|

RESULT 69
Q861K3 PRELIMINARY; PRT; 272 AA.
AC Q861K3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE MHC class I antigen (Fragment).
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=130;
RA Chung C., Leib S.R., Fraser D.G., Ellis S.A., McGuire T.C.;
RT "Sequence-based typing of horse classical MHC class I genes.";
RT Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY176118; AAO18134.1; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG-cl.
DR InterPro; IPR003006; IG_MHC.
DR InterPro; IPR001039; MHC_I.
DR Pfam; PF00047; IG; 1.
DR Pfam; PF00129; MHC_I; 1.
DR PRINTS; PR01638; MHCCLASSI.
DR PRODOM; PD000050; MHC_I; 1.
DR SMART; SM00407; Igc1; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
DR PROSITE; PS00290; IG_MHC; 1.
FT NON_TER 1
FT NON_TER 272
SQ SEQUENCE 272 AA; 30220 MW; F7FC66FFB06558CF CRC64;

Query Match 0.7%; Score 8; DB 7; Length 272;

```

Best Local Similarity 100.0%; Pred. No. 53;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 380 VILGAVGA 387
DB 232 VILGAVGA 239
|||||

RESULT 70
Q8P8NG PRELIMINARY; PRT; 278 AA.
AC Q8P8NG;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)
DE ABC transporter sugar permease.
GN LAGC OR XC2204.
OS Xanthomonas campestris (pv. campestris).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC Xanthomonadaceae; Xanthomonas.
OX NCBI_TaxID=340;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 33913 / NCPPB 528;
MEDLINE=22022145; PubMed=12024217;
RA da Silva A.C.R., Ferro J.A., Reinach P.C., Farah C.S., Furlan L.R.,
RA Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA Canarotte G., Cannavan F., Cardozo J., Chamberg F., Ciapina L.P.,
RA Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA Formighieri B.F., Franco M.C., Greggio C.C., Gruber A.,
RA Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.P.,
RA Locall E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA Moreira L.M., Rossi M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA Pereira H.A., Novo M.T.M., Okura V.K., Silva C., de Souza R.F.,
RA Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA Setubal J.C., Kitajima J.P.;
RP "Comparison of the genomes of two Xanthomonas pathogens with differing
RT host specificities."
RL Nature 417:459-463 (2002).
DR EMBL; AE012328; AAM41484.1; --
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; P:transp.
DR Pfam; PF00528; BPD_transp.
DR Complete proteome.
SQ SEQUENCE 278 AA; 30673 MW; FB45244D924B7361 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 278;
Best Local Similarity 100.0%; Pred. No. 54;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLAL 1159
DB 17 GGLLLAL 24
|||||

RESULT 71
Q8TPV4 PRELIMINARY; PRT; 290 AA.
AC Q8TPV4;
DT 01-JUN-2002 (TRENBLrel. 21, Created)
DT 01-JUN-2002 (TRENBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE Uncharacterized domain specific for M.kandleri, MK-34 family, a.
GN MK1344.
OS Methanopyrus kandleri.
OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
OC Methanopyrus.

```

```

OX NCBI_TaxID=2320;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=AV19 / DSM 6324 / JCM 9639;
MEDLINE=21927647; PubMed=11930014;
RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
RA Shcherbinina O.V., Shakhova V.V., Belova G.I., Aravind L.,
RA Natsale D.A., Rogozin I.B., Tatusov R.L., Wolf Y.I., Stettin K.O.,
RA Malykh A.G., Kochin E.V., Kozlovskiy S.A.;
RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
RT and monophyly of archaeal methanogens."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649 (2002).
DR EMBL; AE010427; AAM02557.1; --
KW Complete proteome.
SQ SEQUENCE 290 AA; 32754 MW; 706E6883976ADF91 CRC64;

Query Match 0.7%; Score 8; DB 17; Length 290;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 268 VITDGESH 275
DB 161 VITDGESH 168
|||||

RESULT 72
Q94K03 PRELIMINARY; PRT; 299 AA.
AC Q94K03;
DT 01-DEC-2001 (TRENBLrel. 19, Created)
DT 01-DEC-2001 (TRENBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Unknown protein.
GN A1G002N01.8.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurasids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF370497; AAK43874.1; --
SQ SEQUENCE 299 AA; 34028 MW; B2AB03C4A184D3E CRC64;

Query Match 0.7%; Score 8; DB 10; Length 299;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLV 1160
DB 44 LLLALLLV 51
|||||

RESULT 73
Q98BD8 PRELIMINARY; PRT; 307 AA.
AC Q98BD8;
DT 01-OCT-2001 (TRENBLrel. 18, Created)
DT 01-OCT-2001 (TRENBLrel. 18, Last sequence update)
DT 01-MAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein mll5620.
GN MLL5620.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;

```

```

RN SEQUENCE FROM N.A.
RC STRAIN=NAFF303059;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA MEDLINE=21082930; PubMed=11214968;
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR ENBL; AP003007; BAB52034.1; -.
DR InterPro; IPR002781; DUF81.
DR Pfam; PF01925; DUF81; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 307 AA; 31951 MW; 938FFCD3BC4AC16 CRC64;

Query Match 0.7%; Score 8; DB 16; Length 307;
Best Local Similarity 100.0%; Pred.No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVLA 1162
DB 275 LLALLVLA 282
|||||

RESULT 74
Q92S85 PRELIMINARY; PRT; 308 AA.
AC Q92S85;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical transmembrane protein SMC02232.
GN R00537 OR SMC02232.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
RA Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ransperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
DR EMBL; AL591784; CAC45109.1; -.
DR InterPro; IPR002781; DUF81.
DR Pfam; PF01925; DUF81; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 308 AA; 32465 MW; CB6497E5E16BC0D CRC64;

Query Match 0.7%; Score 8; DB 16; Length 308;
Best Local Similarity 100.0%; Pred.No. 59;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1155 LLALLVLA 1162
DB 275 LLALLVLA 282
|||||

RESULT 75
Q8CB84 PRELIMINARY; PRT; 323 AA.
ID Q8CB84;
AC Q8CB84;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

```

```

DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Integrin alpha 2 (Fragment).
GN ITGA2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Bone;
RC MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK036572; BAC29485.1; -.
DR MGD; MGI:96600; Itga2.
DR GO; GO:0008305; C:integrin complex; IEA.
DR GO; GO:0004895; F:cell adhesion receptor activity; IEA.
DR GO; GO:0007160; P:cell-matrix adhesion; IEA.
DR InterPro; IPR000413; Integrin_alpha.
DR InterPro; IPR002035; VWF_A.
DR Pfam; PF00092; vwa; 1.
DR PRINTS; PR00453; VWFADOMAIN.
DR SMART; SM00191; Int_alpha; 1.
DR SMART; SM00327; VWA; 1.
DR PROSITE; PS50234; VWFA; 1.
FT NON_TER 323
FT SEQUENCE 323 AA; 35025 MW; 05A4E6A44283724A CRC64;

Query Match 0.7%; Score 8; DB 11; Length 323;
Best Local Similarity 100.0%; Pred.No. 61;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GDVYKCPV 78
DB 75 GDVYKCPV 82
|||||

Search completed: September 21, 2004, 13:41:59
Job time : 78 secs

```

This Page Blank (uspto)

OM protein - protein search, using sw model
Run on: September 21, 2004, 13:05:19 ; Search time 148 Seconds
(without alignments)
2268.016 Million cell updates/sec

Title: US-09-980-403-2
Perfect score: 1188
Sequence: 1 MDLPGLVVAWALS-LWPGFT.....FRSARRRPEGLDTPKYLE 1188

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 1586107 seqs, 282547505 residues

Word size : 0
Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : A_Geneseq_29Jan04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1188	100.0	1188	4	AAB30929 Amino aci
2	928	78.1	1189	4	ABG12949 Novel hum
3	926	77.9	1188	4	AAU14231 Human nov
4	926	77.9	1188	4	AAB50085 Human A25
5	926	77.9	1188	5	AAU10551 Human A25
6	911	76.7	1188	4	AAU14467 Human nov
7	911	76.7	1188	7	ADE09956 Novel pro
8	911	76.7	1189	3	AAB25582 IRGALL pr
9	911	76.7	1189	6	ABR58364 Human nov
10	911	76.7	1189	6	ADA27054 Human nov
11	911	76.7	1189	7	ADE63570 Human Pro
12	911	76.7	1189	8	ADE63570 Human Pro
13	810	68.2	1034	3	AAB25590 Protein e
14	810	68.2	1034	6	ADA27062 Human nov
15	810	68.2	1034	8	ADE66592 Novel hum
16	753	63.4	1120	6	ABR58365 Human NOV
17	489	41.2	707	4	AAU19663 Human nov
18	489	41.2	707	5	ABP47883 Human pol
19	489	41.2	707	7	ADC10845 Human ext
20	193	16.2	193	5	AAU76854 Human int
21	190	16.0	193	5	AAU76863 Human int
22	122	10.3	1188	4	AAB50087 Murine A2
23	122	10.3	1188	5	AAU10552 Murine A2
24	104	8.8	360	7	ADE08585 Novel pro
25	89	7.5	109	5	ABG66673 Human nov

99 8 0.7 17 2 AAW65098 Canine be
100 8 0.7 17 2 AAW72832 Canine al

ALIGNMENTS

RESULT 1

AAB30929
ID AAB30929 standard; protein; 1188 AA.

XX AC AAB30929;

XX DT 02-APR-2001 (first entry)

XX Amino acid sequence of a human alpha11 integrin chain.

XX Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;
XX osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
XX wound healing; trauma; rheumatoid arthritis; osteoarthritis;
XX osteoporosis; cartilage damage; bone damage; cartilage.

XX OS Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..22 /note= "signal peptide"

FT Region 951..972 /note= "972"

FT Domain 1142..1164 /note= "leucine zipper"

FT /note= "transmembrane domain"

XX W0200075187-A1.

XX 14-DEC-2000.

XX 31-MAY-2000; 2000WO-SF001135.

XX 03-JUN-1999; 99SE-00002056.

XX (ACTI-) ACTIVE BIOTECH AB.

XX Gullberg D;

XX WPI; 2001-071061/08.

XX N-PSDB; AAC86871.

XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit
PT alpha 11 in association with subunit beta, useful for treating muscle
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX Disclosure; Fig 2a-c; 79pp; English.

XX The present sequence represents a human integrin subunit, designated
CC alpha11. The alpha11 polynucleotide and polypeptide are useful as markers
CC of cell target molecules, such as fibroblasts, muscle cells,
CC chondrocytes, osteoblasts, mesenchymally derived cells or stem cells.
CC They are also used for determining the differential-stage of cells during
CC differentiation, development in pathological conditions, in tissue
CC regeneration, in transplantation or in therapeutic and physiological
CC repair of tissues. The pathological conditions involving subunit alpha11
CC are selected from damage of cells, muscle dystrophy, fibrosis, wound
CC healing, trauma, rheumatoid arthritis, osteoarthritis and osteoporosis,
CC damage of cartilage and bone, and cartilage and bone diseases. The
CC polypeptide is useful for detecting the formation of cartilage during
CC embryonic development, for detecting physiological therapeutic repair of
CC cartilage and muscle, for selection and analysis, or for sorting,
CC isolating or purification of chondrocytes and muscle cells, for detecting
CC regeneration of cartilage or chondrocytes during transplantation of
CC cartilage or chondrocytes during transplantation of cartilage or
CC chondrocytes, respectively, or of muscle or muscle cells during
CC transplantation of muscle or muscle cells, respectively, and for studies
CC of differentiation of chondrocytes or muscle cells

XX Sequence 1188 AA;

Query Match 100.0%; Score 1188; DB 4; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1188; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVAVWALSLWPGFTDTFNMDTRKPRVTPGSRTPAFGYTVQCHDISGNKWLVGGA 60
DB 1 MDLPRGLVAVWALSLWPGFTDTFNMDTRKPRVTPGSRTPAFGYTVQCHDISGNKWLVGGA 60
QY 61 PLEINGYQKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNMRGLGLSLATNPKDNSFLA 120
DB 61 PLEINGYQKTGDVYKCPVIHGNCTKLNLRVTLSNVSEKDNMRGLGLSLATNPKDNSFLA 120
QY 121 CSPLWSHECGSSYYTTGMCRSVNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSIYPWVE 180
DB 121 CSPLWSHECGSSYYTTGMCRSVNSNFRFSKTVAPALQRCOTYMDIVIVLDGNSIYPWVE 180
QY 181 VQHEFLINILKKFYIGPQIQVGVVQVGEDVVEHFLNDYRSVDVVEAAASHIEQRCGTET 240
DB 181 VQHEFLINILKKFYIGPQIQVGVVQVGEDVVEHFLNDYRSVDVVEAAASHIEQRCGTET 240
QY 241 RTAFGIEFARSEAFQKGRGAKKWMIVITDGSHPDLEKVIQSERDNVTRYAVAVL 300
DB 241 RTAFGIEFARSEAFQKGRGAKKWMIVITDGSHPDLEKVIQSERDNVTRYAVAVL 300
QY 301 GYNRRGINPETFLNETKYIASPDDKHFNVTDEALKDIDVALGDRIFSLEGTNNKET 360
DB 301 GYNRRGINPETFLNETKYIASPDDKHFNVTDEALKDIDVALGDRIFSLEGTNNKET 360
QY 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNWGAVALKETSAGKVIPLRESYLKEFFPEELKN 420
DB 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDNWGAVALKETSAGKVIPLRESYLKEFFPEELKN 420
QY 421 HGAYLGYTVTVSVSSRQGRVYVAGAPRNFHTKVIILFTMHNRSLSLTIHOAMRQOIGSYF 480
DB 421 HGAYLGYTVTVSVSSRQGRVYVAGAPRNFHTKVIILFTMHNRSLSLTIHOAMRQOIGSYF 480
QY 481 GSEITTSVDIDGQVTDVLLVGA PMYNEGRERKGVVYVELQNRVFNCTLKDSSYQNA 540
DB 481 GSEITTSVDIDGQVTDVLLVGA PMYNEGRERKGVVYVELQNRVFNCTLKDSSYQNA 540
QY 541 RFGSSIASVVDLNDQSDYNVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRTASELATG 600
DB 541 RFGSSIASVVDLNDQSDYNVVGAPLEDNHAGAIYIFHGFRGSILKTPKQRTASELATG 600
QY 601 LQVFGCSIHQDLNEDGLDILAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
DB 601 LQVFGCSIHQDLNEDGLDILAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRVTPRAHLDGEGDRFTNRAV 720
DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVIGIRYNATMDERRVTPRAHLDGEGDRFTNRAV 720
QY 721 LLSSGGELCERINFHVLDTADYVVKPTVFSEVSELEDPDHGMDDGWPPTLLSVSPFWNG 780
DB 721 LLSSGGELCERINFHVLDTADYVVKPTVFSEVSELEDPDHGMDDGWPPTLLSVSPFWNG 780
QY 781 CNEDEHCVFDLVDARSDDLPTAMEYQCVLRKPAQDCSAYTSLSFDTTFVIEESTQRVAV 840
DB 781 CNEDEHCVFDLVDARSDDLPTAMEYQCVLRKPAQDCSAYTSLSFDTTFVIEESTQRVAV 840
QY 841 EATLENRGENAVSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQCNVSYPPF 900
DB 841 EATLENRGENAVSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQCNVSYPPF 900
QY 901 RAKAKVAFRLDSEFSKSI FLHHLIELELAAGSDNSNRDSTKEDNVAPLRFLHLYEADVLFT 960
DB 901 RAKAKVAFRLDSEFSKSI FLHHLIELELAAGSDNSNRDSTKEDNVAPLRFLHLYEADVLFT 960
QY 961 RSSLSLHYEVKLNLSLERYDGGPPPPSCIFRQNLGLPFIHGMMKTIPIATRGNNRL 1020

Db 961 RSSLSHYEVKLSRLERYDYGIPFFSCIFRIQNLGLFPIHGMWMTIPIATRSNRL 1020
 QY 1021 KIRDFLTDEANTSCNIGNSTEYRPTVEEDLRAPOLNHNNSDVSNINCIRLVPNQE 1080
 Db 1021 KIRDFLTDEANTSCNIGNSTEYRPTVEEDLRAPOLNHNNSDVSNINCIRLVPNQE 1080
 QY 1081 NFHLGNLWLSLKALKYKSMKIMVNAALQRFHSPFFIFREEDPSRQIEFISKQEDWQV 1140
 Db 1081 NFHLGNLWLSLKALKYKSMKIMVNAALQRFHSPFFIFREEDPSRQIEFISKQEDWQV 1140
 QY 1141 PIIWIVGSTLGLLALLVIALRLKLGPRGARRRREGLDPTPKVLE 1188
 Db 1141 PIIWIVGSTLGLLALLVIALRLKLGPRGARRRREGLDPTPKVLE 1188

RESULT 2
 ID ABG12949 standard; protein; 1189 AA.
 XX ABG12949;

DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #12940.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.

XX Homo sapiens.

XX WO200175067-A2.

XX 11-OCT-2001.

XX 30-MAR-2001; 2001WO-US008631.

XX 31-MAR-2000; 2000US-00540217.

PR 23-AUG-2000; 2000US-00643167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI; 2001-639362/73.

DR N-PSDB; AAS77136.

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.

XX Claim 20; SEQ ID NO 43308; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at

CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1189 AA;
 Query Match 78.1%; Score 928; DB 4; Length 1189;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 1028; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDLPRLGVAWALSLSLWPGFTDTPNMDTRPRVPIPGSRTAFPGYTVQOHDISGNKWLAVGA 60
 Db 1 MDLPRLGVAWALSLSLWPGFTDTPNMDTRPRVPIPGSRTAFPGYTVQOHDISGNKWLAVGA 60
 QY 61 PLETNGVOKTGDVYKCPVHGNCTKLNLRVTLSNVSEKDNRLGLSLATPKNSFLA 120
 Db 61 PLETNGVOKTGDVYKCPVHGNCTKLNLRVTLSNVSEKDNRLGLSLATPKNSFLA 120
 QY 121 CSPLWHECGSSYTTGMCSSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIIPWVE 180
 Db 121 CSPLWHECGSSYTTGMCSSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSIIPWVE 180
 QY 181 VQHFLINILKFKYIGPQIQGVVQVGEDVHVEHFLNDRSVKDVVEAASHIEQGGTET 240
 Db 181 VQHFLINILKFKYIGPQIQGVVQVGEDVHVEHFLNDRSVKDVVEAASHIEQGGTET 240
 QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLEKVIQOSERNVTRYAVAVL 300
 Db 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGHSDSPDLEKVIQOSERNVTRYAVAVL 300
 QY 301 GYNRRGINPETFLNEIKYIASDPDDKHFNVTDEAALKDIVDALGDRIFSLGTKNKT 360
 Db 301 GYNRRGINPETFLNEIKYIASDPDDKHFNVTDEAALKDIVDALGDRIFSLGTKNKT 360
 QY 361 SFGLEMSQTFSSHVVVEDGVLLGAVGAYDNWGLKETSAGKVIPLRESYLKEPPELKN 420
 Db 361 SFGLEMSQTFSSHVVVEDGVLLGAVGAYDNWGLKETSAGKVIPLRESYLKEPPELKN 420
 QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPFNHTGKVIPLTMNNRSLTIHQAMRGOQIGSYF 480
 Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPFNHTGKVIPLTMNNRSLTIHQAMRGOQIGSYF 480
 QY 481 GSEITSDVIDDGVTDVLLVGAPMYFNEGERGKVVYVELRQNFVYNGTLKDSHSYQNA 540
 Db 481 GSEITSDVIDDGVTDVLLVGAPMYFNEGERGKVVYVELRQNFVYNGTLKDSHSYQNA 540
 QY 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFPGSILKTPKQITASELATG 600
 Db 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFPGSILKTPKQITASELATG 600
 QY 601 LQYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 Db 601 LQYFGCSIHGQDLNEDGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDGGRFTNRAV 720
 Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMDERRYTPRAHLDGGRFTNRAV 720
 QY 721 LLSSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGPMDDGHPPTTLRVSPFWNG 780
 Db 721 LLSSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGPMDDGHPPTTLRVSPFWNG 780
 QY 781 CNEDEHCVDPDLVLDARSDDLPTAMEYCORVLRKPAQDCSAYTSLSDTTVFIIESTRQRAV 840
 Db 781 CNEDEHCVDPDLVLDARSDDLPTAMEYCORVLRKPAQDCSAYTSLSDTTVFIIESTRQRAV 840
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLQKESDSDGSEICVNEERLQKQCVNSVPPFF 900
 Db 841 EATLENRGENAYSTVLNISQSANLQFASLQKESDSDGSEICVNEERLQKQCVNSVPPFF 900
 QY 901 RAKAKVAFRLDSEFSKSTFLHLELELAAGSDSNRSDTKEDNVAPLRFHLKYEADVLFT 960
 Db 901 RAKAKVAFRLDSEFSKSTFLHLELELAAGSDSNRSDTKEDNVAPLRFHLKYEADVLFT 960
 QY 961 RSSLSHYEVKLSRLERYDYGIPFFSCIFRIQNLGLFPIHGMWMTIPIATRSNRL 1020

Db 961 RSSLSHYEVKLSLSRYDGIQPPFCIFRIQNLGLFPIHGMWMTIPIATRSNRL 1020
 Qy 1021 KLRDLTDE 1029
 Db 1021 KLRDLTDE 1029
 RESULT 3
 ID AAU14231 standard; protein; 1188 AA.
 AC AAU14231;
 XX 24-OCT-2001 (first entry)
 XX Human novel protein #102.
 XX Human; novel protein; Antianemic; osteopathic; antiinflammatory;
 XX immunomodulatory; cytostatic; neuroprotective; vulnery; nootropic;
 XX anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;
 XX antibacterial; antiallergic; dermatological; haemostatic; antiaesthetic;
 XX thrombolytic; immunogen; antibody; gene therapy; neurological disorder;
 XX Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;
 XX tissue regeneration; immune disorder.
 XX Homo sapiens.
 XX WO200155437-A2.
 XX 25-JAN-2001; 2001WO-US002623.
 XX 25-JAN-2000; 2000US-00491404.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Liu C, Drmanac RT;
 XX WPI; 2001-451939/48.
 XX N-PSDB; AAS22536.
 XX Isolated polypeptides useful for treating anti-inflammatory diseases,
 XX nervous system disorders, and for regenerating bone and cartilage.
 XX Example 4; Page 578-581; 894pp; English.
 XX The invention relates to polynucleotides encoding novel human proteins or
 XX their active domains. The polypeptides, polynucleotides and antibodies
 XX raised against the polypeptides are used in a method of treatment of a
 XX mammal and prevention of disorders caused by the aberrant protein
 XX expression or activity. The polypeptides can be used as molecular weight
 XX markers, food supplements, and in antibody production. The polypeptides
 XX are used to identify compounds which bind to the polypeptides.
 XX Polynucleotides of the invention are used as probes and primers, for
 XX sequencing, for chromosome or gene mapping, in the production of
 XX recombinant proteins, and in generating anti-sense DNA or RNA and in gene
 XX therapy. Polypeptides of the invention can be used to target drugs to a
 XX tumour, in assays to determine biological activity, to raise
 XX antibodies/elicit an immune response, to determine quantitative protein
 XX levels, as tissue markers, and to isolate receptors or ligands.
 XX Polypeptides of the invention may also be useful in treating platelet
 XX disorders, stem cell disorders, regenerating bone, cartilage, tendon,
 XX ligament and/or nerve tissue, wound healing, treating burns, promoting
 XX the proliferation, differentiation and survival of stem cells, as a
 XX contraceptive treating osteoporosis and osteoarthritis, anaemia,
 XX Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
 XX sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
 XX fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
 XX versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
 XX diseases, nervous system disorders, and infection. The present sequence
 XX represents a protein of the invention

XX Sequence 1188 AA;
 SQ
 Query Match 77.9%; Score 926; DB 4; Length 1188;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MDLPRGLVVAWALSILWPGFTDTEFNDTRKPRVPGSTAPFGYTVQOHDISGNKWLVVGA 60
 Db 1 MDLPRGLVVAWALSILWPGFTDTEFNDTRKPRVPGSTAPFGYTVQOHDISGNKWLVVGA 60
 Qy 61 PLETNGYQKTGDVYKCPVIHGNCTKLNGLRVTLNSVSRKDNMRLGLSLATNPKDNSFLA 120
 Db 61 PLETNGYQKTGDVYKCPVIHGNCTKLNGLRVTLNSVSRKDNMRLGLSLATNPKDNSFLA 120
 Qy 121 CSPLWSHECGSSVYTTGMSRVNSNFRFSKTVPALQRCQTYMDIVIVLDGNSIYFWWE 180
 Db 121 CSPLWSHECGSSVYTTGMSRVNSNFRFSKTVPALQRCQTYMDIVIVLDGNSIYFWWE 180
 Qy 181 VQHFLLINILKFFYIGPQIQVGVVQYGEDVVFHFLNDYRSVKDVVEAASHIQRGGTET 240
 Db 181 VQHFLLINILKFFYIGPQIQVGVVQYGEDVVFHFLNDYRSVKDVVEAASHIQRGGTET 240
 Qy 241 RTAFGIEFARSEAPQKGRKGAKKVMITVITDGSHSDSPDLEKVIQOSERDNVTRYAVVL 300
 Db 241 RTAFGIEFARSEAPQKGRKGAKKVMITVITDGSHSDSPDLEKVIQOSERDNVTRYAVVL 300
 Qy 301 GYNRRGINPETFLNEIKYIASPDDKHPNVTDAAALKDIDVALGDRIFSLBGTNNET 360
 Db 301 GYNRRGINPETFLNEIKYIASPDDKHPNVTDAAALKDIDVALGDRIFSLBGTNNET 360
 Qy 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDWNGAVLKETSAGKVIPLRESYLKFEPEELKN 420
 Db 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDWNGAVLKETSAGKVIPLRESYLKFEPEELKN 420
 Qy 421 HGAYLGVTVTSVSSRQGRVYVAGAPRHNHTGVILFTWHNRSFTIHOAMRQQIGSYF 480
 Db 421 HGAYLGVTVTSVSSRQGRVYVAGAPRHNHTGVILFTWHNRSFTIHOAMRQQIGSYF 480
 Qy 481 GSEITSDVDIDGCVTDVLLVGAPMYFNEGRERKVVYVELRQNRVYNGTLDKSHSYQNA 540
 Db 481 GSEITSDVDIDGCVTDVLLVGAPMYFNEGRERKVVYVELRQNRVYNGTLDKSHSYQNA 540
 Qy 541 RFGSSIASVDELNODSVNDVVGAPLEDNHAGNIYIFHGFGRSILKTPQRTASELATG 600
 Db 541 RFGSSIASVDELNODSVNDVVGAPLEDNHAGNIYIFHGFGRSILKTPQRTASELATG 600
 Qy 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 Db 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
 Qy 661 RSRGRATCLAAFLCFTPIFLAPHFQTTTGIIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720
 Db 661 RSRGRATCLAAFLCFTPIFLAPHFQTTTGIIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720
 Qy 721 LLSGGQELCERINFHVLDTADVVKPTFSVEVSLSDPDHGMPLDGMPTTLRVSPFWNG 780
 Db 721 LLSGGQELCERINFHVLDTADVVKPTFSVEVSLSDPDHGMPLDGMPTTLRVSPFWNG 780
 Qy 781 CNEDEHCVPLDLARSDLPTAMEYQVRVLRPAQDCSAYTILSFDTTFFIESTQRVAV 840
 Db 781 CNEDEHCVPLDLARSDLPTAMEYQVRVLRPAQDCSAYTILSFDTTFFIESTQRVAV 840
 Qy 841 EATLENRGENAVSTVLNISQSANLOFASLIQKEDSDGSECVNEERRLQKQVCNVSYPFF 900
 Db 841 EATLENRGENAVSTVLNISQSANLOFASLIQKEDSDGSECVNEERRLQKQVCNVSYPFF 900
 Qy 901 RAKAKVAFRLDSEFSKSIPLHLEIELAAGSDNSERDSTKEDNAPLRFHLKYEADVLFT 960
 Db 901 RAKAKVAFRLDSEFSKSIPLHLEIELAAGSDNSERDSTKEDNAPLRFHLKYEADVLFT 960
 Qy 961 RSSLSHYEVKLSLSRYDGIQPPFCIFRIQNLGLFPIHGMWMTIPIATRSNRL 1020

```

Db 961 RSSLSHVEKPNSSLDYDGGPPFCIFRQNLGLFPIHGMKMTIPATRSNRL 1020
Qy 1021 KLRLDTDEANTSCNIMWNSYRPTPVEEDLRAPOLNHSNSDVSNICNIRLVNQEI 1080
Db 1021 KLRLDTDEANTSCNIMWNSYRPTPVEEDLRAPOLNHSNSDVSNICNIRLVNQEI 1080
Qy 1081 NFHLGNLWLSRLKALKYKSMKWNAALQRFHSPFIFREEDPSROI 1128
Db 1081 NFHLGNLWLSRLKALKYKSMKWNAALQRFHSPFIFREEDPSROI 1128

RESULT 4
ID AAB50085 standard; protein; 1188 AA.
XX AAB50085;
XX 19-MAR-2001 (first entry)
XX Human A259.
XX Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
XX kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
XX rheumatoid arthritis.
XX Homo sapiens.
XX Key
XX Location/Qualifiers
XX 1..1141
XX /label= Extracellular_domain
XX 1..122
XX /label= Signal_peptide
XX 23..1188
XX /label= Mature_protein
XX 39..74
XX /label= Integrin_alpha_subunit_repeat_domain_#1
XX 115..157
XX /label= Integrin_alpha_subunit_repeat_domain_#2
XX 164..345
XX /label= I_domain
XX 367..392
XX /label= Integrin_alpha_subunit_repeat_domain_#3
XX 421..455
XX /label= Integrin_alpha_subunit_repeat_domain_#4
XX 478..516
XX /label= Integrin_alpha_subunit_repeat_domain_#5
XX 540..575
XX /label= Integrin_alpha_subunit_repeat_domain_#6
XX 602..640
XX /label= Integrin_alpha_subunit_repeat_domain_#7
XX 1142..1164
XX /label= Transmembrane_domain
XX 1165..1188
XX /label= Cytoplasmic_domain
XX WO200073339-A1.
XX 07-DEC-2000.
XX 15-MAY-2000; 2000WO-US013262.
XX 28-MAY-1999; 99US-00322790.
XX 27-APR-2000; 2000US-00561263.
XX (MILL-) MILLENNIUM PHARM INC.
XX Pan Y, Lora JM;
XX WPI; 2001-041142/05.
XX N-FSDB; AAC91901, AAC91902.
XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and
XX diagnosis of fibrosis, e.g. of the liver.

```

XX Claim 8; Fig 1; 164pp; English.

XX The present sequence is human integrin alpha subunit, A259. A259 is homologous with the alpha and alpha0 integrin subunits and is overexpressed in fibrosis. A259 is implicated in regulation of cell proliferation, differentiation and/or function of many different cell types. Inhibitors of A259 activity are useful for the treatment of liver disease, particularly fibrosis, and also fibrosis in other organs (specifically lung and kidney). In addition, A259 can be used for treatment and prevention of cancer, osteoporosis, acute myeloid leukaemia, HIV infection, and rheumatoid arthritis

XX Sequence 1188 AA;

Query Match 77.9%; Score 926; DB 4; Length 1188;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 MDLPRGLVAVAWALSLWPGFTDTFNMDTRKPRVTPGSRAPFGVTVQOHDISGNKWLAVGA 60
Db 1 MDLPRGLVAVAWALSLWPGFTDTFNMDTRKPRVTPGSRAPFGVTVQOHDISGNKWLAVGA 60
Qy 61 PLENGYQKTGDVYKCPVHGNCTKLNLRVTLSNYSERKDNRLGLSLATNPKDNSPLA 120
Db 61 PLENGYQKTGDVYKCPVHGNCTKLNLRVTLSNYSERKDNRLGLSLATNPKDNSPLA 120
Qy 121 CSPLWSHECGSSYYTTGMCSSRVNSNRFSTKVPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db 121 CSPLWSHECGSSYYTTGMCSSRVNSNRFSTKVPALQRCQTYMDIVIVLDGNSIYPWVE 180
Qy 181 VQHFELINILKKFYIGPGQIQGVVQYGEDVVFHFLNDYRSKDVVVEAASHIEQRGTT 240
Db 181 VQHFELINILKKFYIGPGQIQGVVQYGEDVVFHFLNDYRSKDVVVEAASHIEQRGTT 240
Qy 241 RTAGIEFARSEAFQKGRGAKKVMIVTDGSHSDSPDLEKVIQOSEKDNVTRYAVVL 300
Db 241 RTAGIEFARSEAFQKGRGAKKVMIVTDGSHSDSPDLEKVIQOSEKDNVTRYAVVL 300
Qy 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTVDEAALKDIVDALGDRIFSLEGTNKN 360
Db 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTVDEAALKDIVDALGDRIFSLEGTNKN 360
Qy 361 SFGLEMSQTFSSHVVEDGVLGAVGAYDNGAVLKTSAGKVIPLRESYLKEFPEELKN 420
Db 361 SFGLEMSQTFSSHVVEDGVLGAVGAYDNGAVLKTSAGKVIPLRESYLKEFPEELKN 420
Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRPNHTGKVIPLTMNNRSLTIHQAMRQQIGSYF 480
Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPRPNHTGKVIPLTMNNRSLTIHQAMRQQIGSYF 480
Qy 481 GSEITSVDIDGQVTVLLVAGAPMYFNEGERKQVYVELRQNRFYNGTLDKSHSYQNA 540
Db 481 GSEITSVDIDGQVTVLLVAGAPMYFNEGERKQVYVELRQNRFYNGTLDKSHSYQNA 540
Qy 541 RFGSSIASVRDLNQDSYNDVVVGAPLEDNHAGAIYFHGFRGSILKTPKQRTASELATG 600
Db 541 RFGSSIASVRDLNQDSYNDVVVGAPLEDNHAGAIYFHGFRGSILKTPKQRTASELATG 600
Qy 601 LQFEGCSIHQGLDNLNEDGLIDLAGVAGNVLWSRPVQINASLHFEPSKINIHRDCK 660
Db 601 LQFEGCSIHQGLDNLNEDGLIDLAGVAGNVLWSRPVQINASLHFEPSKINIHRDCK 660
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRVTPRAHLDEGDRFNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIRYNATMDERRVTPRAHLDEGDRFNRAV 720
Qy 721 LLSSGQELCERINPHVLDIADYKVPVTFSEVSELEDPDHGPMDDGWPTLRSVPFWNG 780
Db 721 LLSSGQELCERINPHVLDIADYKVPVTFSEVSELEDPDHGPMDDGWPTLRSVPFWNG 780
Qy 781 CNEDEHCVPLVDARSDDLPTAMEYQORVLRKPAQCSAYTISFDITVFIIESTRORVAV 840

```

Db 781 CNEDEHCVPLDLVLDARSPLTAMEYQCVLRKPAQDCSAYTSLSDTTVFIIESTRQVAV 840
 Qy 841 EATLENRGENAYSTVLNISQSANIQFASLIQKEDSDGSIQVNEERLQKQVNSYPPF 900
 Db 841 EATLENRGENAYSTVLNISQSANIQFASLIQKEDSDGSIQVNEERLQKQVNSYPPF 900
 Qy 901 RAKAKVAPFLDSERFSKIFLHLHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT 960
 Db 901 RAKAKVAPFLDSERFSKIFLHLHLEIELAAGSDSNERDSTKEDNVAPLRFHLKYEADVLT 960
 Qy 961 RSSLSHYEVKNSLERYDGIQGFSCIFRIQNLGLFPIHGMMKTIPIATRSNRL 1020
 Db 961 RSSLSHYEVKNSLERYDGIQGFSCIFRIQNLGLFPIHGMMKTIPIATRSNRL 1020
 Qy 1021 KLRDFTDEANTSCNMGNSTEYRPTVEEDLRAPOLNHSNDVVSINIRLVPNQEI 1080
 Db 1021 KLRDFTDEANTSCNMGNSTEYRPTVEEDLRAPOLNHSNDVVSINIRLVPNQEI 1080
 Qy 1081 NFHLGNLWLSLKALKYKSKMIMVNAALQRFHSPFIFREDPQRSQI 1128
 Db 1081 NFHLGNLWLSLKALKYKSKMIMVNAALQRFHSPFIFREDPQRSQI 1128

RESULT 5
 AAU10551
 ID AAU10551 standard; protein; 1188 AA.
 XX AC AAU10551;
 XX AC
 DT 14-FEB-2002 (first entry)
 XX
 DE Human A259 polypeptide.
 XX
 KW Human; A259; integrin alpha subunit; integrin alpha 10; secreted protein;
 KW liver disease; fibrosis; lung; kidney; bone associated disorder; blood;
 KW cartilage associated disorder; haematopoietic disorder; bone marrow;
 KW immune related disease; apoptotic disorder; neuronal tissue disease;
 KW neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic;
 KW nephrotropic; immunomodulator; anti-inflammatory; neuroprotective;
 KW antithrombotic; antianemic; antiallergic; antiasthmatic; dermatological;
 KW antidiabetic; anticonvulsant; antiparkinsonian.
 XX
 OS Homo sapiens.
 XX
 PH Location/Qualifiers
 FT 1. .1141
 FT Domain /note= "Extracellular domain"
 FT Peptide 1. .22
 FT /note= "Signal peptide"
 FT Protein 23. .1188
 FT /note= "Mature human A259"
 FT Domain 37. .90
 FT /note= "Integrin alpha repeat domain. The specification
 FT states that this domain exists in human A549"
 FT 115. .157
 FT Domain /note= "Integrin alpha repeat domain"
 FT 164. .345
 FT Domain /note= "I domain or Von Willebrand Factor type A domain"
 FT 367. .392
 FT /note= "Integrin alpha repeat domain"
 FT 421. .472
 FT Domain /note= "Integrin alpha repeat domain. The specification
 FT states that this domain exists in human A549"
 FT 476. .532
 FT Domain /note= "Integrin alpha repeat domain. The specification
 FT states that this domain exists in human A549"
 FT 538. .593
 FT Domain /note= "Integrin alpha repeat domain. The specification
 FT states that this domain exists in human A549"
 FT 600. .654
 FT Domain /note= "Integrin alpha repeat domain. The specification
 FT states that this domain exists in human A549"
 FT 1142. .1164

FT Domain /note= "Transmembrane domain"
 FT 1165. .1188
 XX /note= "Cytoplasmic domain"
 PN W0200181414-A2.
 XX 01-NOV-2001.
 XX 27-APR-2001; 2001WO-US013516.
 XX 27-APR-2000; 2000US-00561263.
 PR (MILL-) MILLENNIUM PHARM INC.
 XX Pan Y, Lora J;
 PI WPI; 2002-041397/05.
 DR N-PSDB; AAS16873.
 XX
 PT New A259 nucleic acids and polypeptides, which comprise integrin alpha
 PT subunit, useful for diagnosing, preventing or treating e.g. liver
 PT disease, kidney or lung fibrosis, cancers, blood disorders or immune
 PT related diseases.
 XX
 PS Claim 9; Fig 1; 168pp; English.
 XX
 CC The invention relates to human and murine A259 nucleic acid molecules
 CC which encode secreted proteins with homology to integrin alpha subunits,
 CC specifically to integrin alpha 10. The A259 polypeptide and nucleic acid
 CC are useful for treating liver disease or fibrosis, particularly kidney
 CC fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also
 CC useful for diagnosing, preventing or treating cartilage and bone
 CC associated disorders (such as bone cancer, achondroplasia, myeloma,
 CC fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and
 CC acute myeloid leukaemia, haemophilia, anaemia and thalassaemia), immune
 CC related diseases (such as HIV, viral infections, cancers, T cell
 CC autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g.
 CC asthma and psoriasis), apoptotic disorders (such as systemic lupus
 CC erythematosus and insulin-dependent diabetes mellitus), diseases of the
 CC neuronal tissues (such as epilepsy and muscular dystrophy) and
 CC neurodegenerative diseases (such as Parkinson's disease and Huntington's
 CC disease). This sequence represents the human A259 polypeptide
 XX
 SQ Sequence 1188 AA;
 Query Match 77.9%; Score 926; DB 5; Length 1188;
 Best Local Similarity 99.8%; Pred. No. 0;
 Matches 1126, Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MDLPRLGLVAVALSLLWPGFTDTFNMDTRKPRVPGSRTAFFGYTVQOQHDISGNKMLVGA 60
 Db 1 MDLPRLGLVAVALSLLWPGFTDTFNMDTRKPRVPGSRTAFFGYTVQOQHDISGNKMLVGA 60
 Qy 61 PLETNGVQKTGDYVKCPVTHGKCTKLNLRVTLSNVSEKDNVRLGLSLATPKNSFLA 120
 Db 61 PLETNGVQKTGDYVKCPVTHGKCTKLNLRVTLSNVSEKDNVRLGLSLATPKNSFLA 120
 Qy 121 CSPLWSHECGSSYTTTGMCSRNSNFRFSKTVAPALQRCQTYMDIVILVDGNSIYPWVE 180
 Db 121 CSPLWSHECGSSYTTTGMCSRNSNFRFSKTVAPALQRCQTYMDIVILVDGNSIYPWVE 180
 Qy 181 VQHFNLINILKFFIGPQIQVGVQVQGEDYVHFEHLNDYRSKDVVEAASHTEQGGTET 240
 Db 181 VQHFNLINILKFFIGPQIQVGVQVQGEDYVHFEHLNDYRSKDVVEAASHTEQGGTET 240
 Qy 241 RTAFGEFARSEAFQKGRKGAKVMIVITDGHSDSPLEKVIQOESRDNVTRAVAVL 300
 Db 241 RTAFGEFARSEAFQKGRKGAKVMIVITDGHSDSPLEKVIQOESRDNVTRAVAVL 300
 Qy 301 GYNRRGINPETFLNBIKIYIASDPDDKHFFNVTDAAALKDIVDALGDRIFSLEGTNKNET 360
 Db 301 GYNRRGINPETFLNBIKIYIASDPDDKHFFNVTDAAALKDIVDALGDRIFSLEGTNKNET 360

QY	361	SFGLMSQTGFSSHVVYEDGVLLGAVGAYDWNGAVLKETSGAKVPIPIRESYLKEFPBELXN	420
Db	361	SFGLMSQTGFSSHVVYEDGVLLGAVGAYDWNGAVLKETSGAKVPIPIRESYLKEFPBELXN	420
QY	421	HGAYLGTYTTSVYSSRQGVVYAGAPRFNHTGVILLFTWHNRESLTIHQAMRQQQLGSYF	480
Db	421	HGAYLGTYTTSVYSSRQGVVYAGAPRFNHTGVILLFTWHNRESLTIHQAMRQQQLGSYF	480
QY	481	GSEITSVDIDGQVTDVLLVGAPMYNEGRERKGVVYELRQNRFYNGTLKDSHSHYQNA	540
Db	481	GSEITSVDIDGQVTDVLLVGAPMYNEGRERKGVVYELRQNRFYNGTLKDSHSHYQNA	540
QY	541	RGSSSTASVRDLNODSYNDVVYCAPLEDNHAGAIYIFHGFPGSILTKPKQRIATASELATG	600
Db	541	RGSSSTASVRDLNODSYNDVVYCAPLEDNHAGAIYIFHGFPGSILTKPKQRIATASELATG	600
QY	601	LOVFGSGIHQGLDNLNEDGLDIDAVGALGNAILWSRPVQVINASLHFEPSKINIFHRDCK	660
Db	601	LOVFGSGIHQGLDNLNEDGLDIDAVGALGNAILWSRPVQVINASLHFEPSKINIFHRDCK	660
QY	661	RSGRDATCLAAFCFTPIFIAPHFQTTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAV	720
Db	661	RSGRDATCLAAFCFTPIFIAPHFQTTVGIRYNATMDERRYTPRAHLDEGGDRFTNRAV	720
QY	721	LLSSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGEMLDGQWPTTLRVSVPFWNG	780
Db	721	LLSSGQELCERINFHVLDTADYVKPVTFSVEYSLEDPDHGEMLDGQWPTTLRVSVPFWNG	780
QY	781	CNEDEHCVPDLVLDARSDLPTAMEYCORVLKPAQDCSAYTLSDFTTVFIIESTRQVAV	840
Db	781	CNEDEHCVPDLVLDARSDLPTAMEYCORVLKPAQDCSAYTLSDFTTVFIIESTRQVAV	840
QY	841	EATLENRGENAYGTVLNISOSANLOFASLIQKEDSDGSIECVNEERLQKQVCNVSYPPF	900
Db	841	EATLENRGENAYGTVLNISOSANLOFASLIQKEDSDGSIECVNEERLQKQVCNVSYPPF	900
QY	901	RAKAKVAFRLDSEFSKSI FLHLHLEILAAGSDNERDSTKEDNVAPLRFLKYEADVLF	960
Db	901	RAKAKVAFRLDSEFSKSI FLHLHLEILAAGSDNERDSTKEDNVAPLRFLKYEADVLF	960
QY	961	RSSSLSHYEVKLNSSLERYDGIQPPFSCIFRIQNLGLFPPIHGMWMTIPIATRSGNRL	1020
Db	961	RSSSLSHYEVKPNSSLERYDGIQPPFSCIFRIQNLGLFPPIHGMWMTIPIATRSGNRL	1020
QY	1021	KLADFDTDEANTSCINWGNSTERYPTPEVEDLRAPQLNHSNSDVWSINCINRLVPNQEI	1080
Db	1021	KLADFDTDEANTSCINWGNSTERYPTPEVEDLRAPQLNHSNSDVWSINCINRLVPNQEI	1080
QY	1081	NFHLGLNLWLRSLKALKYKSNKIMVNAALORQPHSPFIREDPSROI 1128	
Db	1081	NFHLGLNLWLRSLKALKYKSNKIMVNAALORQPHSPFIREDPSROI 1128	
RESULT 6			
AAU14467			
XX	ID	AAU14467 standard; protein; 1188 AA.	
XX	AC	AAU14467;	
XX	DT	24-OCT-2001 (first entry)	
XX	DE	Human novel protein #338.	
XX	KW	Human; novel protein; Antianaemic; osteopathic; antiinflammatory;	
XX	KW	immunomodulatory; cytostatic; neuroprotective; vulnery; nootropic;	
XX	KW	anticonvulsant; antiarthritic; cerebroprotective; antifungal; antiviral;	
XX	KW	antibacterial; antiallergic; dermatological; haemostatic; antiasthmatic;	
XX	KW	thrombolytic; immunogen; antibody; gene therapy; neurological disorder;	
XX	KW	Parkinson's disease; inflammatory disorder; cancer; asthma; osteoporosis;	
XX	KW	tissue regeneration; immune disorder.	
XX	XX	Homo sapiens.	

XX	PN	WO200155437-A2.
XX	PD	02-AUG-2001.
XX	PF	25-JAN-2001; 2001WO-US002623.
XX	PR	25-JAN-2000; 2000US-00491404.
XX	PA	(HYSE-) HYSSEQ INC.
XX	PI	Tang YT, Liu C, Drmanac RT;
XX	DR	WPI; 2001-451939/48.
XX	DR	N-PSDB; AAS22772.
XX	PT	Isolated polypeptides useful for treating anti-inflammatory diseases,
XX	PT	nervous system disorders, and for regenerating bone and cartilage.
XX	PP	Example 4; Page 828-831; 894pp; English.
XX	CC	The invention relates to polynucleotides encoding novel human proteins or
XX	CC	their active domains. The polypeptides, polynucleotides and antibodies
XX	CC	raised against the polypeptides are used in a method of treatment of a
XX	CC	mammal and prevention of disorders caused by the aberrant protein
XX	CC	expression or activity. The polypeptides can be used as molecular weight
XX	CC	markers, food supplements, and in antibody production. The polypeptides
XX	CC	are used to identify compounds which bind to the polypeptides.
XX	CC	Polynucleotides of the invention are used as probes and primers, for
XX	CC	sequencing, for chromosome or gene mapping, in the production of
XX	CC	recombinant proteins, and in generating anti-sense DNA or RNA and in gene
XX	CC	therapy. Polypeptides of the invention can be used to target drugs to a
XX	CC	tumour, in assays to determine biological activity, to raise
XX	CC	antibodies/elicit an immune response, to determine quantitative protein
XX	CC	levels, as tissue markers, and to isolate receptors or ligands.
XX	CC	Polypeptides of the invention may also be useful in treating platelet
XX	CC	disorders, stem cell disorders, regenerating bone, cartilage, tendon,
XX	CC	ligament and/or nerve tissue, wound healing, treating burns, promoting
XX	CC	the proliferation, differentiation and survival of stem cells, as a
XX	CC	contractile, treating osteoporosis and osteoarthritis, anaemia, anaemia,
XX	CC	Alzheimer's, Parkinson's and Huntington's diseases, amyotrophic lateral
XX	CC	sclerosis, stroke, immune deficiencies resulting from bacterial, viral or
XX	CC	fungal infection or from autoimmunity, cancer, allergy, asthma, graft-
XX	CC	-versus-host disease, eczema, haemophilia, thrombosis, anti-inflammatory
XX	CC	diseases, nervous system disorders, and infection. The present sequence
XX	CC	represents a protein of the invention
XX	SQ	Sequence 1188 AA;
		Query Match 76.7%; Score 911; DB 4; Length 1188;
		Best Local Similarity 100.0%; Pred. No. 0;
		Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0
QY	1	MDLPRGLVAVALSILWPQGTTFENMTRKPRVIPGSRFAFGYTQQOHDISGNKLVLVGA 60
DB	1	MDLPRGLVAVALSILWPQGTTFENMTRKPRVIPGSRFAFGYTQQOHDISGNKLVLVGA 60
QY	61	PLEINGYOKTGDVVKCPVIHGNCITKLNIGRVTLSNVSERKNMRGLSLATNPXDNFSFLA 120
DB	61	PLEINGYOKTGDVVKCPVIHGNCITKLNIGRVTLSNVSERKNMRGLSLATNPXDNFSFLA 120
QY	121	CSPLMWSHCSSGYTTGCMSRVNSNFRSKTVAPALQCQTVMIVILDGSNSLYPWVE 180
DB	121	CSPLMWSHCSSGYTTGCMSRVNSNFRSKTVAPALQCQTVMIVILDGSNSLYPWVE 180
QY	181	VQHFLINTLKIFYIGPGQIQGVVOYGEDVHVEFHLNDYRSVKDYVEAAASHIEQRGGTET 240
DB	181	VQHFLINTLKIFYIGPGQIQGVVOYGEDVHVEFHLNDYRSVKDYVEAAASHIEQRGGTET 240
QY	241	RTAFGIETARSAEPFKGGRKGAKKMYIITGESHSDPDLKVIQCSREDNVTRAVAVL 300
DB	241	RTAFGIETARSAEPFKGGRKGAKKMYIITGESHSDPDLKVIQCSREDNVTRAVAVL 300

```
Query Match          76.7%; Score 911; DB 4; Length 1188;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911: Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	1	MDLPRLVVAVALS	LPQGTDTFNM	DTKRKRV	IPGSR	TAFPGY	TVOQHD	ISGNKML	VGVA	60		
Db	1	MDLPRLVVAVALS	LPQGTDTFNM	DTKRKRV	IPGSR	TAFPGY	TVOQHD	ISGNKML	VGVA	60		
Qy	61	PLETNGYQKTG	DVVKCP	VIHGNCT	KNLGRV	TLSNYSER	KDNMR	LGSLATNP	KNSFLA	120		
Db	61	PLETNGYQKTG	DVVKCP	VIHGNCT	KNLGRV	TLSNYSER	KDNMR	LGSLATNP	KNSFLA	120		
Qy	121	CSPLWSH	ECGSGSY	TTGMC	SVNSVNS	FRFSK	TVAPAL	QRCTYMD	IVIVLDG	SNSIYPWVE	180	
Db	121	CSPLWSH	ECGSGSY	TTGMC	SVNSVNS	FRFSK	TVAPAL	QRCTYMD	IVIVLDG	SNSIYPWVE	180	
Qy	181	VQHFLIN	ILTKFY	IPGGQ	IQGVVQY	GEDVVH	FHFLND	YRSVKDV	VEAAASH	IEQRGTET	240	
Db	181	VQHFLIN	ILTKFY	IPGGQ	IQGVVQY	GEDVVH	FHFLND	YRSVKDV	VEAAASH	IEQRGTET	240	
Qy	241	RTAFG	IEFARS	BAFQ	KGKRG	AKKVM	IVITD	GESHSD	PDLEK	VIQSERDNT	RYAVAVL	300
Db	241	RTAFG	IEFARS	BAFQ	KGKRG	AKKVM	IVITD	GESHSD	PDLEK	VIQSERDNT	RYAVAVL	300

QY 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTDAAALDKDIDVALGDRIFSLGTTNKNET 360
 DB 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTDAAALDKDIDVALGDRIFSLGTTNKNET 360
 QY 361 SFGLSEMSQTFSSHWEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420
 DB 361 SFGLSEMSQTFSSHWEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420
 QY 421 HGAYLGVTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQQQIGSYF 480
 DB 421 HGAYLGVTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQQQIGSYF 480
 QY 481 GSEITSDIDGDGTVDLLVGAPMYFNEGRERGKVVYVELQRNFRVYNGTLKDSHSYQNA 540
 DB 481 GSEITSDIDGDGTVDLLVGAPMYFNEGRERGKVVYVELQRNFRVYNGTLKDSHSYQNA 540
 QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGILKTPKQITASELATG 600
 DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGILKTPKQITASELATG 600
 QY 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPKINIHRDCK 660
 DB 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPKINIHRDCK 660
 QY 661 RSRGDATCLAAFLCFTPIFLAPHFQITTVGIRYNATWDBERRYTPRAHLDGEGDRFTNRAV 720
 DB 661 RSRGDATCLAAFLCFTPIFLAPHFQITTVGIRYNATWDBERRYTPRAHLDGEGDRFTNRAV 720
 QY 721 LLSGQELCERINPHVLDADYKVPVTFSSVEYSLEDDPHGPMDDGWPITLRSVVPFWNG 780
 DB 721 LLSGQELCERINPHVLDADYKVPVTFSSVEYSLEDDPHGPMDDGWPITLRSVVPFWNG 780
 QY 781 CNEDEHCVPLVDARSDLPTAMEYQCVLRKPAQDCSAYTILSFDTTVIIESTRQAV 840
 DB 781 CNEDEHCVPLVDARSDLPTAMEYQCVLRKPAQDCSAYTILSFDTTVIIESTRQAV 840
 QY 841 EATLENRGENAYTVLNISANLQFASLIQKEDSDGSEICVNEERLQKQCNVSPFF 900
 DB 841 EATLENRGENAYTVLNISANLQFASLIQKEDSDGSEICVNEERLQKQCNVSPFF 900
 QY 901 RAKAKVAFRLD 911
 DB 901 RAKAKVAFRLD 911

RESULT 7
ADE09956

ID ADE09956 standard; protein; 1188 AA.

AC ADE09956;

XX ADE09956;

DT 29-JAN-2004 (first entry)

DE Novel protein-related contig polypeptide sequence #544.

KW novel gene; novel protein; tissue marker; molecular weight marker;
 KW chromosome marker; Genetic disorder; contig.

XX Unidentified.

XX WO2003054152-A2.

XX 03-JUL-2003.

XX 10-DEC-2002; 2002MO-US039555.

XX 10-DEC-2001; 2001US-0339739P.

XX 11-DEC-2001; 2001US-0339453P.

XX 14-MAR-2002; 2002US-0365091P.

XX 14-MAR-2002; 2002US-0365384P.

XX 12-APR-2002; 2002US-0372381P.

XX 12-APR-2002; 2002US-0372615P.

XX 22-APR-2002; 2002US-00128558.

PR 24-APR-2002; 2002US-0376045P.
 XX (HYSE-) HYSEQ INC.
 XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
 PI Ghosh M, Xue AJ, Wehrman T, Wang G, Zhou P, Drmanac RI, Wang Z;
 PI Ma Y, Wang D, Chen R, Xu C, Boyie BJ;
 XX WPI; 2003-569235/53.
 XX New polynucleotides, useful for expressing recombinant proteins for
 PT analysis, characterization or therapeutic use, or as markers for tissues
 PT in which the corresponding protein is preferentially expressed.
 XX Disclosure; SEQ ID NO 3022; 1177pp; English.
 XX The invention comprises the amino acid and coding sequences of novel
 CC proteins. The DNA and protein sequences of the invention are useful as:
 CC markers for tissues in which the corresponding protein is preferentially
 CC expressed; as molecular weight markers on gels; as chromosome markers or
 CC tags; to identify chromosomes or to map related gene positions; and to
 CC compare with endogenous DNA sequences in patients to identify potential
 CC genetic disorders. The present amino acid sequence was used in the
 CC exemplification of the invention.
 XX Sequence 1188 AA;
 SQ Query Match 76.7%; Score 911; DB 7; Length 1188;
 Best Local Similarity 100.0%; Pred. NO. 0;
 Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MDLPRGLVWAWLSLWPGFTDTFNMDTRKPRVTPGRTAFPGYVVOQHDLSGNKWLAVGA 60
 DB 1 MDLPRGLVWAWLSLWPGFTDTFNMDTRKPRVTPGRTAFPGYVVOQHDLSGNKWLAVGA 60
 QY 61 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSERKDNMRGLSLATNPKNSFLA 120
 DB 61 PLETNGYOKTGDVYKCPVIHGNCTKLNLRVTLSNVSERKDNMRGLSLATNPKNSFLA 120
 QY 121 CSPLWSHECCSSVYTTGMCSSRVNSNFRSKTAPALORCOTYMDIVILDGNSIYFWVE 180
 DB 121 CSPLWSHECCSSVYTTGMCSSRVNSNFRSKTAPALORCOTYMDIVILDGNSIYFWVE 180
 QY 181 VQHFLINILKKFYIGPGQIQVGVVQYGEDVWFHFLNDYRSVKDVVEAASHIEQRGTET 240
 DB 181 VQHFLINILKKFYIGPGQIQVGVVQYGEDVWFHFLNDYRSVKDVVEAASHIEQRGTET 240
 QY 241 RTAFGIEFARSEAFQGGKGAKKWIMIVITDGSHPDLEKVIQOESRDNTVRYAVVL 300
 DB 241 RTAFGIEFARSEAFQGGKGAKKWIMIVITDGSHPDLEKVIQOESRDNTVRYAVVL 300
 QY 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTDAAALDKDIDVALGDRIFSLGTTNKNET 360
 DB 301 GYNRRGINPETFLNEIKYIASDDDDKHFFNVTDAAALDKDIDVALGDRIFSLGTTNKNET 360
 QY 361 SFGLSEMSQTFSSHWEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420
 DB 361 SFGLSEMSQTFSSHWEDGVLGAVGAYDNGAVLKETSAGKVIPLRESYLKEFPPELKN 420
 QY 421 HGAYLGVTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQQQIGSYF 480
 DB 421 HGAYLGVTVTSVSSRQGRVYVAGAPFNHTGKVLFTMNNRSLTIHQAMRQQQIGSYF 480
 QY 481 GSEITSDIDGDGTVDLLVGAPMYFNEGRERGKVVYVELQRNFRVYNGTLKDSHSYQNA 540
 DB 481 GSEITSDIDGDGTVDLLVGAPMYFNEGRERGKVVYVELQRNFRVYNGTLKDSHSYQNA 540
 QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGILKTPKQITASELATG 600
 DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFSGILKTPKQITASELATG 600
 QY 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPKINIHRDCK 660
 DB 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPKINIHRDCK 660

601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFPSPKINIFHRDCK 660
661 RSGRDATCLAAFLCPTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720
661 RSGRDATCLAAFLCPTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720
721 LLSGQELCERINFHVLDTADVVPKPTFVSVEYSLEDPDHPGMLDGMPTTLRVSPVPFWNG 780
721 LLSGQELCERINFHVLDTADVVPKPTFVSVEYSLEDPDHPGMLDGMPTTLRVSPVPFWNG 780
781 CNEDEHCVPLVDLARSPLTAMEYQORVLRPAQDCSAYTILSFDTTTFPIESTQRVAV 840
781 CNEDEHCVPLVDLARSPLTAMEYQORVLRPAQDCSAYTILSFDTTTFPIESTQRVAV 840
841 EATLENRGENAYSTVLNISQSANLQPASLIQKEDSDGSECNERRLQKQVCNVSYPFF 900
841 EATLENRGENAYSTVLNISQSANLQPASLIQKEDSDGSECNERRLQKQVCNVSYPFF 900
901 RAKAKVAFRLD 911
901 RAKAKVAFRLD 911
RESULT 8
ID AAB25582
AC AAB25582 standard; protein; 1189 AA.
XX
XX
XX
XX 21-NOV-2000 (first entry)
XX
XX ITGA11 protein encoded by human secreted protein gene #7.
DE
XX Secreted protein; immunosuppressant; anti-inflammatory; antithrombotic;
XX antithrombotic, dermatological; antiproliferative; antitumor; antithrombotic;
XX anticancer; vulvar; antiviral; antibacterial; antifungal;
XX immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;
XX multiple sclerosis; inflammatory disorder; inflammatory bowel disease;
XX Crohn's disease; nephritis; hyperproliferative disorder;
XX cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;
XX melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.
XX
XX Homo sapiens.
XX
XX WO200029435-A1.
XX
XX 25-MAY-2000.
XX
XX 27-OCT-1999; 99WO-US025031.
XX
XX 28-OCT-1998; 98US-0105971P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;
XX Greene JM;
XX
XX WPI; 2000-387742/33.
XX
XX N-PSDS; AAA80612.
XX
XX Isolated nucleic acid molecules encoding human secreted proteins are used
XX for the prevention, amelioration and treatment of autoimmune,
XX inflammatory, hyperproliferative and cardiovascular disorders, cancer,
XX wounds, and infectious diseases.
XX
XX Claim 1; Fig 19A-F; 803pp; English.
XX
XX The present invention relates to 12 secreted human proteins and the
XX nucleotide sequences encoding them. The polynucleotide sequences given in
XX AAA80612-A80623 encode the 12 secreted protein sequences given in
XX AAB25576-B25593. The human secreted proteins have various activities
XX dependent on the tissues in which they are expressed. Examples of the
XX activities of the proteins include: immunosuppressant; anti-inflammatory;

antiarthritic; antirheumatic, dermatological; antiproliferative;
XX antiarteriosclerotic; anticancer; vulvar; antiviral; antibacterial;
XX and antifungal activity. The proteins, polypeptides, agonists and
XX antagonists may be used to treat prevent and/or diagnose various disease,
XX disorders and conditions examples of which include: immune disorders e.g.
XX Addison's disease, rheumatoid arthritis, dermatitis, and multiple
XX sclerosis; inflammatory disorders e.g. inflammatory bowel disease,
XX Crohn's disease and nephritis; hyperproliferative disorders such as
XX arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The
XX proteins and polynucleotide sequences may also be used in wound healing
XX and the treatment of infectious diseases. The human secreted protein gene
XX #7 and protein sequences are represented in sequences AAA80612 and
XX AAB25582. Secreted protein gene #7 is located at position chromosome 15
XX q22.3-23. Sequences AAA80652-A80661 represent genes which are related to
XX the secreted protein gene#7
XX
XX Sequence 1189 AA;
Query Match 76.7%; Score 911; DB 3; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDLPRGLVAVWALSPLPGFTDTFNMDTRKPRVPGSRTAFFGTVQOHDISGNKWLVVGA 60
DB 1 MDLPRGLVAVWALSPLPGFTDTFNMDTRKPRVPGSRTAFFGTVQOHDISGNKWLVVGA 60
QY 61 PLENGYQKTDVYKCPVHGNCTKLNGLRVYLSNVSEKDNMRLGLSLATNPKNDSFLA 120
DB 61 PLENGYQKTDVYKCPVHGNCTKLNGLRVYLSNVSEKDNMRLGLSLATNPKNDSFLA 120
QY 121 CSPLWSHECCSSYTTGCSRVSNSFRFSKTVAPALQRCOTYMDIVLVLDGNSIYPWVE 180
DB 121 CSPLWSHECCSSYTTGCSRVSNSFRFSKTVAPALQRCOTYMDIVLVLDGNSIYPWVE 180
QY 181 VQHEFLNLKFFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDWEAASHIEQGGTET 240
DB 181 VQHEFLNLKFFYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDWEAASHIEQGGTET 240
QY 241 RTAFGIFFARSEAFQKGRKGAKKMIVITDGSHSDPDLKVKVQOERDNTVRYAVL 300
DB 241 RTAFGIFFARSEAFQKGRKGAKKMIVITDGSHSDPDLKVKVQOERDNTVRYAVL 300
QY 301 GYNRRGINPETELNEIKYIASDDPDKHFFNVNTDEAALKDIVDALGDRIFSLGNTKNET 360
DB 301 GYNRRGINPETELNEIKYIASDDPDKHFFNVNTDEAALKDIVDALGDRIFSLGNTKNET 360
QY 361 SFGLEMSQTFSSHVVEDGVLGAVGYDMNGAVLKTSAGKVIPLRESYLKEPPELKN 420
DB 361 SFGLEMSQTFSSHVVEDGVLGAVGYDMNGAVLKTSAGKVIPLRESYLKEPPELKN 420
QY 421 HGAVILGYTVTSVSSRQGRVYVAGAPRHNHTGKVLFTMHNRSITLHQAARGQIGSYF 480
DB 421 HGAVILGYTVTSVSSRQGRVYVAGAPRHNHTGKVLFTMHNRSITLHQAARGQIGSYF 480
QY 481 GSEITSVDIDGVTDLVLLGAPMYFNEGREGRGVYVYELQNRFPVYNGTLKDSHSYQNA 540
DB 481 GSEITSVDIDGVTDLVLLGAPMYFNEGREGRGVYVYELQNRFPVYNGTLKDSHSYQNA 540
QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAYIFHGFSGSLKTPKQITASELATG 600
DB 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAYIFHGFSGSLKTPKQITASELATG 600
QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFPSPKINIFHRDCK 660
DB 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILWSRPVQVNASLHFPSPKINIFHRDCK 660
QY 661 RSGRDATCLAAFLCPTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720
DB 661 RSGRDATCLAAFLCPTPIFLAPHFQTTVIGIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720
QY 721 LLSGQELCERINFHVLDTADVVPKPTFVSVEYSLEDPDHPGMLDGMPTTLRVSPVPFWNG 780
DB 721 LLSGQELCERINFHVLDTADVVPKPTFVSVEYSLEDPDHPGMLDGMPTTLRVSPVPFWNG 780

Db 721 LLSSGQELCERINFHVLDTADYVKPVTFVSVEYSLEDDPHGMLDDGWPTTLRVSVFWMNG 780
 Qy 781 CNEDEHCVPLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSDFTTTFVFIESTRQVAV 840
 Db 781 CNEDEHCVPLVLDARSDLPTAMEYCORVLRKPAQDCSAYTLSDFTTTFVFIESTRQVAV 840
 Qy 841 EATLENGENAYSTVLNISOSANLPASLIQKEDSDGSIIECVNEERLQKQVCNVSYPFF 900
 Db 841 EATLENGENAYSTVLNISOSANLPASLIQKEDSDGSIIECVNEERLQKQVCNVSYPFF 900
 Qy 901 RAKAKVAFRLD 911
 Db 901 RAKAKVAFRLD 911

RESULT 9

ABR58364
 ID ABR58364 standard; protein; 1189 AA.

XX AC ABR58364;
 XX DT 07-JUL-2003 (first entry)
 XX DE Human NOV2a.
 XX KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;
 KW antiparkinsonian; antilipaeamic; gene therapy; metabolic disorder;
 KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;
 KW neurodegenerative disorder; Alzheimer's disease; immune disorder;
 KW haematopoietic disorder.
 OS Homo sapiens.
 XX PN W02003029423-A2.
 XX PD 10-APR-2003.

XX PF 02-OCT-2002; 2002WO-US031358.

XX PR 02-OCT-2001; 2001US-0326483P.

XX PR 05-OCT-2001; 2001US-0327342P.

XX PR 09-OCT-2001; 2001US-0327917P.

XX PR 09-OCT-2001; 2001US-0328029P.

XX PR 09-OCT-2001; 2001US-0328044P.

XX PR 09-OCT-2001; 2001US-0328056P.

XX PR 15-OCT-2001; 2001US-0328849P.

XX PR 17-OCT-2001; 2001US-0330142P.

XX PR 22-OCT-2001; 2001US-0341058P.

XX PR 24-OCT-2001; 2001US-0339266P.

XX PR 24-OCT-2001; 2001US-0343629P.

XX PR 29-OCT-2001; 2001US-0349575P.

XX PR 01-NOV-2001; 2001US-0346357P.

XX PR 12-APR-2002; 2002US-0371972P.

XX PR 12-APR-2002; 2002US-0371980P.

XX PR 17-APR-2002; 2002US-0373261P.

XX PR 19-APR-2002; 2002US-0373805P.

XX PR 23-APR-2002; 2002US-0374738P.

XX PR 16-MAY-2002; 2002US-0381101P.

XX PR 17-MAY-2002; 2002US-0381635P.

XX PR 29-MAY-2002; 2002US-0383830P.

XX PR 01-OCT-2002; 2002US-00262839.

XX (CURA-) CURAGEN CORP.

XX PA Alsbrook JP, Anderson DW, Boldog FL, Burgess CE, Catterton E;

XX PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W;

XX PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK;

XX PI Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ;

XX PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M;

XX WPI; 2003-381625/36.

DR N-PSDB; ACC72076.

XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
 PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
 PT dyslipidemia, and in chromosome mapping, tissue typing or
 PT pharmacogenomics.

XX Claim 1; Page 105; 487pp; English.

XX The present invention relates to novel human NOV proteins and their
 CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
 CC proteins are useful in manufacturing a medicament for treating a syndrome
 CC associated with a human disease. The NOV proteins and coding sequences
 CC may be used to diagnose, treat or prevent metabolic disorders such as
 CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
 CC disorders such as Alzheimer's disease or Parkinson's disease, immune
 CC disorders, haematopoietic disorders and various dyslipidaemias

SQ Sequence 1189 AA;

Query Match 76.7%; Score 911; DB 6; Length 1189;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLPRGLVVAWALS LWPFGTDTFNMDTRKPRVPGSGRTAFFGYTVQOHDISGNKWL VVGA 60

Db 1 MDLPRGLVVAWALS LWPFGTDTFNMDTRKPRVPGSGRTAFFGYTVQOHDISGNKWL VVGA 60

Qy 61 PLENGYQKTDGVYKCPVIHGNCTKMLGRVTLNVSRKDNRLGLSLATNPKDNSFLA 120

Db 61 PLENGYQKTDGVYKCPVIHGNCTKMLGRVTLNVSRKDNRLGLSLATNPKDNSFLA 120

Qy 121 CSPLWSHECGSSYYTTGMC SRVNSNFRFSKTVAPALQRCQTYMDIVILVDGNSIYPWVE 180

Db 121 CSPLWSHECGSSYYTTGMC SRVNSNFRFSKTVAPALQRCQTYMDIVILVDGNSIYPWVE 180

Qy 181 VQHFLINILKFKYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVVEAAASHIEQRCGTET 240

Db 181 VQHFLINILKFKYIGPGQIQGVVQYGEDVVEHFLNDYRSVKDVVEAAASHIEQRCGTET 240

Qy 241 RTAFGLIEFARSEAFQKGRGAKKWMIVITDGESHDSPLLEKVIQOESRDNTVYAVVL 300

Db 241 RTAFGLIEFARSEAFQKGRGAKKWMIVITDGESHDSPLLEKVIQOESRDNTVYAVVL 300

Qy 301 GYNNRGINPETELNEIKYIASDPDDKHFFNVVTDEAALXDIVDALGDRIFSLEGTNNKET 360

Db 301 GYNNRGINPETELNEIKYIASDPDDKHFFNVVTDEAALXDIVDALGDRIFSLEGTNNKET 360

Qy 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDWNGAVLKETSAGKVIPLRESYLKEFPPELQN 420

Db 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDWNGAVLKETSAGKVIPLRESYLKEFPPELQN 420

Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHOAMRQQQIGSYF 480

Db 421 HGAYLGYTVTSVSSRQGRVYVAGAPRFNHTGKVLFTMNNRSLTIHOAMRQQQIGSYF 480

Qy 481 GSEITSDVDIGDGVTVLLVAGPMYNEGREKGVYVYBELQRNFVYNGTLDKSHSYQNA 540

Db 481 GSEITSDVDIGDGVTVLLVAGPMYNEGREKGVYVYBELQRNFVYNGTLDKSHSYQNA 540

Qy 541 RFGSSIASVRDLNQDSYNDVVGAPEDNHAGAIYIFHGFGRSILKTPQRITASELATG 600

Db 541 RFGSSIASVRDLNQDSYNDVVGAPEDNHAGAIYIFHGFGRSILKTPQRITASELATG 600

Qy 601 LQYFGCSIHQGLDNLNEDGLIDLAVGALGNVILWSPVQVINASLHFPEPSKINIFHRDCK 660

Db 601 LQYFGCSIHQGLDNLNEDGLIDLAVGALGNVILWSPVQVINASLHFPEPSKINIFHRDCK 660

Qy 661 RSGRDATCLAAFLCFTPIFLAPHFOTTTGIRYNATMDERRYPTRAHLDEGGDRNTNRAV 720

Db 661 RSGRDATCLAAFLCFTPIFLAPHFOTTTGIRYNATMDERRYPTRAHLDEGGDRNTNRAV 720

Qy 721 LLSSGQELCERINFHVLDTADYVKPVTFVSVEYSLEDDPHGMLDDGWPTTLRVSVFWMNG 780

Db 721 LLSGGQELCERINFHVLDTADYVKPTVTSFVSEYSLSDPDHGPMLDDGWPTTLVSPFPWNG 780
QY 781 CNEDEHCVDPDLVLDARSDLPAMAYCORVLRKPAQDCSAYTLSFDTTFVFIESTQRVAV 840
Db 781 CNEDEHCVDPDLVLDARSDLPAMAYCORVLRKPAQDCSAYTLSFDTTFVFIESTQRVAV 840
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERRRQKQVCNVSYPPF 900
Db 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERRRQKQVCNVSYPPF 900
QY 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911

RESULT 10

ADA27054
ID ADA27054 standard; protein; 1189 AA.

XX AC ADA27054;

XX DT 20-NOV-2003 (first entry)

XX DE Human novel secreted protein from cDNA HOHBY69 #1.

XX KW cytostatic; antiinflammatory; immunomodulator; neuroprotective;

XX KW hemostatic; gene therapy; cancer; inflammation; immune disorder;

XX KW neurological disorder; blood clotting disorder; food additive;

XX KW preservative; human; secreted protein.

XX OS Homo sapiens.

XX PN US2003055231-A1.

XX PD 20-MAR-2003.

XX PR 29-OCT-2001; 2001US-00984130.

XX PR 28-OCT-1998; 98US-0105971P.

XX PR 27-OCT-1999; 99WO-US025031.

XX PR 19-APR-2000; 2000US-0198407P.

XX PR 30-OCT-2000; 2000US-0243792P.

XX PR 18-APR-2001; 2001US-00836353.

XX PA (NIJJ/) NI J.

XX PA (YOUN/) YOUNG P E.

XX PA (KENN/) KENNY J J.

XX PA (OLSE/) OLSEN H S.

XX PA (MOOR/) MOORE P A.

XX PA (WEIY/) WEI Y.

XX PA (GREE/) GREENE J M.

XX PA (RUBE/) RUBEN S M.

XX PA (LIUD/) LIU D.

XX PA (CROC/) CROCKER P R.

XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;

PI Ruben SM, Liu D, Crocker PR;

XX WPT: 2003-567103/53.

DR N-PSDB; ADA27036.

XX New human secreted nucleic acid molecules and polypeptides, useful for

PT preventing, treating, or ameliorating a medical condition, such as

PT cancer, inflammation, immune disorders, neurological and blood clotting

PT disorders.

XX Claim 11; Fig 19; 454pp; English.

XX The invention relates to an isolated nucleic molecule that is at least

CC 95% identical to 18 human cDNA sequences representing 12 novel genes

CC encoding secreted proteins or a polynucleotide fragment of the cDNA

CC sequence contained in American Type Culture Collection (ATCC) deposit No.

CC defined in the specification, its species homologue, a variant or allelic
CC variant of the polynucleotide having a polynucleotide capable of
CC hybridising under conditions the polynucleotide, where the polynucleotide
CC does not hybridise under stringent conditions to a nucleic acid molecule
CC having a nucleotide sequence of only A or T residues. Also included are
CC recombinant vectors, host cells (for producing the polypeptide), the
CC secreted polypeptide (comprising a sequence that is at least 95%
CC identical to a polypeptide fragment, domain, epitope, full-length
CC protein, variant, allelic variant or species homologue), antibodies that
CC specifically bind to the polypeptides, diagnosing, treating, preventing
CC or ameliorating a medical condition by administering the polynucleotide
CC or the polypeptide, the gene corresponding to the cDNA sequence and
CC identifying an activity in a biological assay (by expressing the cDNA
CC sequence in a cell, isolating the supernatant, and detecting an activity
CC in a biological assay and identifying the protein in the supernatant
CC having the activity). The polypeptides, nucleic acids and antibodies are
CC useful for diagnosing a pathological condition or a susceptibility to a
CC pathological condition, for preventing, treating, or ameliorating a
CC medical condition, such as cancer, inflammation and other immune
CC disorders, neurological and blood clotting disorders (many examples are
CC given in the specification). The nucleic acids are also useful for
CC chromosome identification, radiation hybrid mapping or long-range
CC restriction mapping. The polypeptides and antibodies are useful for
CC providing immunological probes for differential identification of the
CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,
CC agonist or antagonist may also be used as a food additive or preservative
CC to increase or decrease storage capabilities, fat content or other
CC nutritional components. The present is a secreted protein of the
CC invention.

XX SQ Sequence 1189 AA;

Query Match 76.7%; Score 911; DB 6; Length 1189;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDLPRGLVVAWALSILWPGFTTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLVGA 60

Db 1 MDLPRGLVVAWALSILWPGFTTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLVGA 60

QY 61 PLETNVYOKTGDVVKPCVHGNCTKLNLRVTLNSVSRKDNMRGLSLAIPKDNLSFLA 120

Db 61 PLETNVYOKTGDVVKPCVHGNCTKLNLRVTLNSVSRKDNMRGLSLAIPKDNLSFLA 120

QY 121 CSPLWSHECGSSYYTGMCSRNSNFRFSKTVA PALQRCTYMDIIVLDGNSIYPWVE 180

Db 121 CSPLWSHECGSSYYTGMCSRNSNFRFSKTVA PALQRCTYMDIIVLDGNSIYPWVE 180

QY 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVVEHEHNDYRSVKDVEAASHIEORGGTET 240

Db 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVVEHEHNDYRSVKDVEAASHIEORGGTET 240

QY 241 RTAFGIEFARGEAFQGGKRGAKKVMIVITDGEHSDSPDLEKVIQOQSERDNNVRYAVAVL 300

Db 241 RTAFGIEFARGEAFQGGKRGAKKVMIVITDGEHSDSPDLEKVIQOQSERDNNVRYAVAVL 300

QY 301 GYNERGINPETFLNEIKYIASDDPDKHFFNVNTEAALKDIVDGLDIFSLGKTNKNET 360

Db 301 GYNERGINPETFLNEIKYIASDDPDKHFFNVNTEAALKDIVDGLDIFSLGKTNKNET 360

QY 361 SFGLEMSOTGRSSHVVVEDGVLGAVYDNGAVLKETSAGKVIPLRESYLKEFPPEELKN 420

Db 361 SFGLEMSOTGRSSHVVVEDGVLGAVYDNGAVLKETSAGKVIPLRESYLKEFPPEELKN 420

QY 421 HGAYLGYTVTSVSSRQGRVVVAGAPRNFHTGKVLFTMHNRSLTTHQAMRGGQIGSYF 480

Db 421 HGAYLGYTVTSVSSRQGRVVVAGAPRNFHTGKVLFTMHNRSLTTHQAMRGGQIGSYF 480

QY 481 GSEITSVDIDGGVTDVLLVGAAPMYFNEGRGKVVYVELQNRVYVNGTLKDSHSYQNA 540

Db 481 GSEITSVDIDGGVTDVLLVGAAPMYFNEGRGKVVYVELQNRVYVNGTLKDSHSYQNA 540

QY 541 RFGSSIASVRDLNQLSDVNDVVVGVVGFLEDNHAIAIYIFHFRGSILTKPKQRTASELATG 600

Db 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFSGILKTPKQITASELATG 600
Qy 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
Db 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660
Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTPRAHLDEGGDFNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTPRAHLDEGGDFNRAV 720
Qy 721 LLSGQELCERINPHVLDADYKVPVTFSEYSLDHPDHPMLDDGWPTTLRVSVFPWNG 780
Db 721 LLSGQELCERINPHVLDADYKVPVTFSEYSLDHPDHPMLDDGWPTTLRVSVFPWNG 780
Qy 781 CNEDEHCVPLDLDARSDLPTAMEYCORVLRKPAQDCSAYTLSFDTVFIESTRORVAV 840
Db 781 CNEDEHCVPLDLDARSDLPTAMEYCORVLRKPAQDCSAYTLSFDTVFIESTRORVAV 840
Qy 841 EATLENRGENAYSTVLNISOSANLQFASLIQKEDSDGSEICVNEERRLOKQVCNVSYPFF 900
Db 841 EATLENRGENAYSTVLNISOSANLQFASLIQKEDSDGSEICVNEERRLOKQVCNVSYPFF 900
Qy 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911

RESULT 11

AD63570
ID ADE63570 standard; protein; 1189 AA.

XX ADE63570;

Dt 29-JAN-2004 (first entry)

DE Human Protein Q9URX5, SEQ ID NO 9514.

XX Human; pain; neuronal tissue; gene therapy;

KW Spinal segmental nerve injury; chronic constriction injury; CCI;

KW Spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WC2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

XX (GEO) GEN HOSPITAL CORP.

PA (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

PT New composition comprising two or more isolated polypeptides, useful for
Pt preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent

CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1189 AA;

Query Match 76.7%; Score 911; DB 7; Length 1189;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MDLPRGLVAVWALS LWPFGFTDTFNMDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60

Db 1 MDLPRGLVAVWALS LWPFGFTDTFNMDTRKPRVPGSRPTAFPGYTVQOHDISGNKWLAVGA 60

Qy 61 PLETNGYQKTGDYVKCPVHGNCTKLNGLRVTLNVSRKDNRLGLSLATNPKDNSFLA 120

Db 61 PLETNGYQKTGDYVKCPVHGNCTKLNGLRVTLNVSRKDNRLGLSLATNPKDNSFLA 120

Qy 121 CSPLSWSEHCSSYYTTGMCSCRVSNSNFRSKTVAPALQRCOTYMDIVLVDGNSIYPWVE 180

Db 121 CSPLSWSEHCSSYYTTGMCSCRVSNSNFRSKTVAPALQRCOTYMDIVLVDGNSIYPWVE 180

Qy 181 VQFLINILKKFYIGPQIQVGVVQYGEDVVFHFLNDYRSVDVVEAASHIEQRGGTET 240

Db 181 VQFLINILKKFYIGPQIQVGVVQYGEDVVFHFLNDYRSVDVVEAASHIEQRGGTET 240

Qy 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGHSDSPDLKVIQOSERDNTVRYAVVL 300

Db 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGHSDSPDLKVIQOSERDNTVRYAVVL 300

Qy 301 GYNNRGINPETFLNEIKYIASPDDKHPFNVTDEAALKDIDVALGDRIFSLEGTNKNET 360

Db 301 GYNNRGINPETFLNEIKYIASPDDKHPFNVTDEAALKDIDVALGDRIFSLEGTNKNET 360

Qy 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEFPEELKN 420

Db 361 SFGLEMSQTGFSSHVVEDGVLLGAVGAYDMNGAVLKETSAGKVIPLRESYLKEFPEELKN 420

Qy 421 HGAYLGYTVTTSVVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRLSLTHQAMRQQQIGSYF 480

Db 421 HGAYLGYTVTTSVVSSRQGRVYVAGAPRFNHTGKVIILFTMHNRLSLTHQAMRQQQIGSYF 480

Qy 481 GSITSDVDIGDGVTDVLLVGAPMYNEGERGKVVYELRQNRFPVNGTLKDSHYQNA 540

Db 481 GSITSDVDIGDGVTDVLLVGAPMYNEGERGKVVYELRQNRFPVNGTLKDSHYQNA 540

Qy 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFSGILKTPKQITASELATG 600

Db 541 RFGSSIASVRDLNQDSYNDVVVVGAPLEDNHAGAIYIFHGFSGILKTPKQITASELATG 600

Qy 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660

Db 601 LOYFGCSIHGOLDLNEGLDLAVGALGNVILWSRPVQINASLHFEPSKINIFHRDCK 660

Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTPRAHLDEGGDFNRAV 720

Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTPRAHLDEGGDFNRAV 720

Db 661 RSGRATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720
Qy 721 LSSGOELCERINFHVLDTADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVFPFNG 780
Db 721 LSSGOELCERINFHVLDTADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVFPFNG 780
Qy 781 CNEDEHCVDPDLVDARSDLPATAMEYQORVLRKPAQDCSAYTLSFTTTFIESTRQAV 840
Db 781 CNEDEHCVDPDLVDARSDLPATAMEYQORVLRKPAQDCSAYTLSFTTTFIESTRQAV 840
Qy 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900
Db 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900
Qy 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911

RESULT 12
ADE86584
ID ADE86584 standard; protein; 1189 AA.
AC ADE86584;
XX
DT 29-JAN-2004 (first entry)
DE Novel human secreted protein #7.
KW human; secreted protein; cancer; liver disorder; hepatitis;
KW neural disorder; Alzheimer's disease.
OS Homo sapiens.
XX
XX US2003129685-A1.
XX 10-JUL-2003.
XX 18-APR-2001; 2001US-00836353.
XX
XX 28-OCT-1998; 98US-0105971P.
XX 27-OCT-1999; 99WO-US025031.
XX 19-APR-2000; 2000US-0198407P.
XX
XX (NLIJ/) NI J.
XX (YOUN/) YOUNG P E.
XX (KENN/) KENNY J J.
XX (OLSE/) OLSEN H S.
XX (MOOR/) MOORE P A.
XX (WEIY/) WEI Y.
XX (GREE/) GREENE J M.
XX (RUBE/) RUBEN S M.
XX
XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
XX Ruben SM;
XX
XX WPI: 2004-020335/02.
XX N-PSDB; ADE86566.
XX
XX New nucleic acid molecule, useful for preparing a medicament for
XX preventing, treating or ameliorating a medical condition e.g. cancer,
XX liver disorders or neural disorders.
XX
XX Claim 11; SEQ ID NO 35; 380pp; English.

XX The invention relates to an isolated nucleic acid sequence, or its
XX allelic variant, a fragment of the cDNA sequence, or its fragment,
XX domain, epitope or species homologue. The nucleic acid is useful for
XX preparing a medicament for preventing, treating or ameliorating a medical
XX condition e.g., cancer, liver disorders such as hepatitis or neural
XX disorders such as Alzheimer's disease. The present sequence represents
XX the amino acid sequence of a novel human secreted protein.

SQ Sequence 1189 AA;
Query Match 76.7%; Score 911; DB 8; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MDLPRLGVAVAWLSLWPGFTDFTNMDTRKPRVTPGSRTPAFQYVTVQOHDISGNKWLWVGA 60
Db 1 MDLPRLGVAVAWLSLWPGFTDFTNMDTRKPRVTPGSRTPAFQYVTVQOHDISGNKWLWVGA 60
Qy 61 PLEFNGYQKTGDDVVKCVIHNCTKLNLRVTLNSVSRKDNMRLGLSLATNPKDSFLA 120
Db 61 PLEFNGYQKTGDDVVKCVIHNCTKLNLRVTLNSVSRKDNMRLGLSLATNPKDSFLA 120
Qy 121 CSPLWSHECGSSYYTTCMCSRVSNSFRFSKTVAPALQRCQTYMDIVVLGDSNSIYPWVE 180
Db 121 CSPLWSHECGSSYYTTCMCSRVSNSFRFSKTVAPALQRCQTYMDIVVLGDSNSIYPWVE 180
Qy 181 VQHELINILKKFYIGPGQIQGVVQYGEDVVFHFLNDYRSKDVVZAAASHIEQGGTET 240
Db 181 VQHELINILKKFYIGPGQIQGVVQYGEDVVFHFLNDYRSKDVVZAAASHIEQGGTET 240
Qy 241 RTAFGIEFARSEAFQKGRKGAKVMITVITDGSNHSPLDLEKVIQSERDNVTRYAVL 300
Db 241 RTAFGIEFARSEAFQKGRKGAKVMITVITDGSNHSPLDLEKVIQSERDNVTRYAVL 300
Qy 301 GYNNRGINPETFLNEIKYIASDDPKHFFNVTDAAALKDIVDALGDRIFSLGTTKNET 360
Db 301 GYNNRGINPETFLNEIKYIASDDPKHFFNVTDAAALKDIVDALGDRIFSLGTTKNET 360
Qy 361 SFGLEMSQTFSSHVWEDGVLGAVGAYDNWAGVLKETSAGKVIPLRESYLKEPPELKN 420
Db 361 SFGLEMSQTFSSHVWEDGVLGAVGAYDNWAGVLKETSAGKVIPLRESYLKEPPELKN 420
Qy 421 HGAYLGYTVTSVSSRQGRVTVAGAPFNHTGKVLFTMHNRSLLTHQMEGQOIGSYF 480
Db 421 HGAYLGYTVTSVSSRQGRVTVAGAPFNHTGKVLFTMHNRSLLTHQMEGQOIGSYF 480
Qy 481 GSEITSDVIDGDGVTDLVLLVGAPMYFNEGRERGVVYVYELQNFVYNGTLKDSHSYQNA 540
Db 481 GSEITSDVIDGDGVTDLVLLVGAPMYFNEGRERGVVYVYELQNFVYNGTLKDSHSYQNA 540
Qy 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFPGSILKTPKQITASELATG 600
Db 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGFPGSILKTPKQITASELATG 600
Qy 601 LQYFGCSIHGOLDLNEGLIDLAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 660
Db 601 LQYFGCSIHGOLDLNEGLIDLAVGALGNVILWSRPVQINASLHPEPSKINIFHRDCK 660
Qy 661 RSGRATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720
Db 661 RSGRATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYPTRAHLDEGGDRFTNRAV 720
Qy 721 LSSGOELCERINFHVLDTADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVFPFNG 780
Db 721 LSSGOELCERINFHVLDTADYVKPVTFSVEYSLEDDPHGPMDDGWPTTLRVSVFPFNG 780
Qy 781 CNEDEHCVDPDLVDARSDLPATAMEYQORVLRKPAQDCSAYTLSFTTTFIESTRQAV 840
Db 781 CNEDEHCVDPDLVDARSDLPATAMEYQORVLRKPAQDCSAYTLSFTTTFIESTRQAV 840
Qy 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900
Db 841 EATLENGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERLQKQVCNVSYPFF 900
Qy 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911

RESULT 13
AAB25590

ID XX AAB25590 standard; protein; 1034 AA.
 XX AC AAB25590;
 XX DT 21-NOV-2000 (first entry)
 XX DE Protein encoded by human secreted protein gene #7 clone HOBY69.
 XX KW Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;
 KW antirheumatic, dermatological; antiproliferative; antiarteriosclerotic;
 KW anticancer; vulnary; antiviral; antibacterial; antifungal;
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;
 KW Crohn's disease; nephritis; hyperproliferative disorder;
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;
 KW melanoma; lymphoma; wound healing; human; chromosome 15 q22.3-23.
 XX OS Homo sapiens.
 XX PN W0200029435-A1.
 XX PD 25-MAY-2000.
 XX PF 27-OCT-1999; 99WO-US025031.
 XX PR 28-OCT-1998; 98US-0105971P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;
 PI Greene JM;
 XX WP1; 2000-387742/33.
 XX PT Isolated nucleic acid molecules encoding human secreted proteins are used
 PT for the prevention, amelioration and treatment of autoimmune,
 PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,
 PT wounds, and infectious diseases.
 XX PS Claim 1; Page 678-682; 803pp; English.
 XX CC The present invention relates to 12 secreted human proteins and the
 CC nucleotide sequences encoding them. The polynucleotide sequences given in
 CC AAA80606-A80623 encode the 12 secreted protein sequences given in
 CC AAB25576-B25593. The human secreted proteins have various activities
 CC dependent on the tissues in which they are expressed. Examples of the
 CC activities of the proteins include: immunosuppressant; anti-inflammatory;
 CC antiarthritic; antirheumatic, dermatological; antiproliferative;
 CC antiarteriosclerotic; anticancer; vulnary; antiviral; antibacterial;
 CC and antifungal activity. The proteins, polypeptides, agonists and
 CC antagonists may be used to treat prevent and/or diagnose various disease,
 CC disorders and conditions examples of which include: immune disorders e.g.
 CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple
 CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,
 CC Crohn's disease and nephritis; hyperproliferative disorders such as
 CC paraneoplasias and purpura; cardiovascular disorders e.g. coronary
 CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The
 CC proteins and polynucleotide sequences may also be used in wound healing
 CC and the treatment of infectious diseases. The human secreted protein gene
 CC #7 and protein sequences are represented in sequences AAA80612 and
 CC AAB25582. Secreted protein gene #7 is located at position chromosome 15
 CC q22.3-23. Sequences AAA80652-A80661 represent genes which are related to
 CC the secreted protein gene #7
 XX SQ Sequence 1034 AA;
 XX Query Match 68.2%; Score 810; DB 3; Length 1034;
 XX Best Local Similarity 99.9%; Pred. NO. 0;
 XX Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDLPRGLVVAWALSLWPGFTDTFNMDTRKPRVPGSTAFFGYTQQHDSGNKWLWVGA 60
 DB 1 MDLPRGLVVAWALSLWPGFTDTFNMDTRKPRVPGSTAFFGYTQQHDSGNKWLWVGA 60

QY 61 PLENGYQKTGDVYKCPVIHGNCTKLNLRVTLNSVSRKDNMRGLGLSLATNPKDNSFLA 120
 DB 61 PLENGYQKTGDVYKCPVIHGNCTKLNLRVTLNSVSRKDNMRGLGLSLATNPKDNSFLA 120
 QY 121 CSPLWSHECGSSYYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIVTLVDCGNSIYPWE 180
 DB 121 CSPLWSHECGSSYYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIVTLVDCGNSIYPWE 180
 QY 191 VOHFLINILKKFYIGPGQIQGVVQYGEDVVEHFLNDYVSVDVVEAAASHIQRGGTET 240
 DB 191 VOHFLINILKKFYIGPGQIQGVVQYGEDVVEHFLNDYVSVDVVEAAASHIQRGGTET 240
 QY 241 RTAFGIEFARSEAFQKGRGAKVMIVITDGSHSDSPDLKVIQCSERDNVTRYAVAVL 300
 DB 241 RTAFGIEFARSEAFQKGRGAKVMIVITDGSHSDSPDLKVIQCSERDNVTRYAVAVL 300
 QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGTHKNET 360
 DB 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGTHKNET 360
 QY 361 SFGLEMSQTFSSHVEDGVLLGAVGAYDNGAVLAKETSAGKVIPLRESYLKEFFPEELKN 420
 DB 361 SFGLEMSQTFSSHVEDGVLLGAVGAYDNGAVLAKETSAGKVIPLRESYLKEFFPEELKN 420
 QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVIILFTMNNRSLTIHQAMRQQIGSYF 480
 DB 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTGKVIILFTMNNRSLTIHQAMRQQIGSYF 480
 QY 481 GSEITTSVIDIDGVDVLLVAGAPMYFNEGRERKVVYVELRQNFYVNGTLKDSHSYQNA 540
 DB 481 GSEITTSVIDIDGVDVLLVAGAPMYFNEGRERKVVYVELRQNFYVNGTLKDSHSYQNA 540
 QY 541 RFSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKORITASELATG 600
 DB 541 RFSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKORITASELATG 600
 QY 601 LOYFGCSIHQGLDINEDGLDLAGVGNALVILWSRPVQINASLHFEPSKINIFHRDCK 660
 DB 601 LOYFGCSIHQGLDINEDGLDLAGVGNALVILWSRPVQINASLHFEPSKINIFHRDCK 660
 QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
 DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMDERRYTTPRAHLDGEGDRFTNRAV 720
 QY 721 LLSGQELCERINPHVLDADYKPVTFVSVEYSELEDDHGPMLDDGWPTTLRVSVPFWNG 780
 DB 721 LLSGQELCERINPHVLDADYKPVTFVSVEYSELEDDHGPMLDDGWPTTLRVSVPFWNG 780
 QY 781 CNEDEHCVPLDLDARSDLPATMEYQCVLRKPAQDCSAYTILSFDTTVFIESTRORVAV 840
 DB 781 CNEDEHCVPLDLDARSDLPATMEYQCVLRKPAQDCSAYTILSFDTTVFIESTRORVAV 840
 QY 841 EATLENGENAYSTVLNISQANLQFASLIQKEDSDGSECVNEERRLOKQVCNVSYPFF 900
 DB 841 EATLENGENAYSTVLNISQANLQFASLIQKEDSDGSECVNEERRLOKQVCNVSYPFF 900
 QY 901 RAKAKVAFRLD 911
 DB 901 RAKAKVAFRLD 911
 RESULT 14
 ADA27062
 ID ADA27062 standard; protein; 1034 AA.
 XX ADA27062;
 AC ADA27062;
 XX 20-NOV-2003 (first entry)
 DT Human novel secreted protein from cDNA HOBY69 #2.
 DE cytostatic; antiinflammatory; immunomodulator; neuroprotective;
 XX

RESULT 15
 ADE86592
 ID ADE86592 standard; protein; 1034 AA.
 XX
 AC ADE86592;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Novel human secreted protein #15.
 XX
 KW human; secreted protein; cancer; liver disorder; hepatitis;
 KW neural disorder; Alzheimer's disease.
 XX
 OS Homo sapiens.
 XX
 FN US2003129685-A1.
 XX
 PD 10-JUL-2003.
 XX
 PF 18-APR-2001; 2001US-00836353.
 XX
 PR 28-OCT-1998; 98US-0105971P.
 PR 27-OCT-1999; 99WO-US025031.
 PR 19-APR-2000; 2000US-0198407P.
 XX
 PA (NIJJ/) NI J.
 PA (YOUN/) YOUNG P E.
 PA (KENN/) KENNY J J.
 PA (OLSE/) OLSEN H S.
 PA (MOOR/) MOORE P A.
 PA (WEIY/) WEI Y.
 PA (GREE/) GREENE J M.
 PA (RUBE/) RUBEN S M.
 XX
 PI Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
 PI Ruben SM;
 XX
 WPI; 2004-020335/02.
 DR N-PSDB; ADE86574.
 XX
 PT New nucleic acid molecule, useful for preparing a medicament for
 PT preventing, treating or ameliorating a medical condition e.g. cancer,
 PT liver disorders or neural disorders.
 XX
 PS Claim 11; SEQ ID NO 43; 380pp; English.
 XX
 CC The invention relates to an isolated nucleic acid sequence, or its
 CC allelic variant, a fragment of the cDNA sequence, or its fragment,
 CC domain, epitope or species homologue. The nucleic acid is useful for
 CC preparing a medicament for preventing, treating or ameliorating a medical
 CC condition e.g., cancer, liver disorders such as hepatitis or neural
 CC disorders such as Alzheimer's disease. The present sequence represents
 CC the amino acid sequence of a novel human secreted protein.
 XX
 SQ Sequence 1034 AA;
 Query Match 68.2%; Score 810; DB 8; Length 1034;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 910; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MDLPRLVAVWALSLEPGFTDFNMTRKPRVTPGSRFAFGYVQOHDISGNKHLVYGA 60
 DB 1 MDLPRLVAVWALSLEPGFTDFNMTRKPRVTPGSRFAFGYVQOHDISGNKHLVYGA 60
 QY 61 PLETNGYQKTGDYKCPVHGNCTKLNGLRVLSNVSEKKNRGLSLATNPKDNSFLA 120
 DB 61 PLETNGYQKTGDYKCPVHGNCTKLNGLRVLSNVSEKKNRGLSLATNPKDNSFLA 120
 QY 121 CSPLASHCEGSSYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVILVDCNSIYPWE 180
 DB 121 CSPLASHCEGSSYTTGMCSSRVNSFRFSKTVAPALQRCQTYMDIVILVDCNSIYPWE 180

QY 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLNDYRSVKDVFAASHIEQGGTET 240
 DB 181 VOHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLNDYRSVKDVFAASHIEQGGTET 240
 QY 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGSHSDSPDLKVIQOQSERDNTRYAVAVL 300
 DB 241 RTAFGIEFARSEAFQKGRKGAKKVMIVITDGSHSDSPDLKVIQOQSERDNTRYAVAVL 300
 QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIVDALGDRIFSLGNTKNET 360
 DB 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIVDALGDRIFSLGNTKNET 360
 QY 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYKLPPEELKN 420
 DB 361 SFGLEMSQTGFSSHVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYKLPPEELKN 420
 QY 421 HGAYLGYTVTSVSSROGRVYVAGAPRFNHTGKVLFTMHNRSLLTHQAMRQQQIGSYF 480
 DB 421 HGAYLGYTVTSVSSROGRVYVAGAPRFNHTGKVLFTMHNRSLLTHQAMRQQQIGSYF 480
 QY 481 GSEITSDVIDGQVTDVLLVAGAPMYFNEGRERGVYVYELRQNRVYNGTLKDSHSYQNA 540
 DB 481 GSEITSDVIDGQVTDVLLVAGAPMYFNEGRERGVYVYELRQNRVYNGTLKDSHSYQNA 540
 QY 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIPHGFRGSLTKTPKORITASELATG 600
 DB 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIPHGFRGSLTKTPKORITASELATG 600
 QY 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQVQINASLHFEPSKINIFHRDCK 660
 DB 601 LQYFGCSIHGQDLNEDGLIDLAVGALGNVILMSRPVQVQINASLHFEPSKINIFHRDCK 660
 QY 661 RSGRDATCLAAFLCFTPIFLAPHQTTVIGIRYNATMDERRYTPRAHLDEGGDRFTNRVAV 720
 DB 661 RSGRDATCLAAFLCFTPIFLAPHQTTVIGIRYNATMDERRYTPRAHLDEGGDRFTNRVAV 720
 QY 721 LLSSGQELCERINPHVLDTADYVKPVTFSVEYSLEDPDHGPMDDGMPPTTLRVSVFPWNG 780
 DB 721 LLSSGQELCERINPHVLDTADYVKPVTFSVEYSLEDPDHGPMDDGMPPTTLRVSVFPWNG 780
 QY 781 CNEDEHCVPLVLDARSDDLPTAMEYQORVLRKPAQCSAYTLSEDTTTFIESTRQAVAV 840
 DB 781 CNEDEHCVPLVLDARSDDLPTAMEYQORVLRKPAQCSAYTLSEDTTTFIESTRQAVAV 840
 QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRLOKQVCNVSYPFF 900
 DB 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSIECVNEERRLOKQVCNVSYPFF 900
 QY 901 RAKAKVAFRLD 911
 DB 901 RAKAKVAFRLD 911
 RESULT 16
 ABR58365
 ID ABR58365 standard; protein; 1120 AA.
 XX
 AC ABR58365;
 XX
 DT 07-JUL-2003 (first entry)
 XX
 DE Human NOV2b.
 XX
 KW Human; NOV; antidiabetic; anorectic; antibacterial; virucide;
 KW immunomodulator; cytostatic; nootropic; neuroprotective; dyslipidaemia;
 KW antiparkinsonian; antihypertensive; gene therapy; metabolic disorder;
 KW diabetes; obesity; infection; cachexia; cancer; Parkinson's disease;
 KW neurodegenerative disorder; Alzheimer's disease; immune disorder;
 KW haematopoietic disorder.
 XX
 OS Homo sapiens.
 XX
 FN WO2003029423-A2.

XX PD 10-APR-2003.
XX PF 02-OCT-2002; 2002WO-US031358.
XX
PR 02-OCT-2001; 2001US-0326483P.
PR 05-OCT-2001; 2001US-0327342P.
PR 03-OCT-2001; 2001US-0327917P.
PR 03-OCT-2001; 2001US-0328029P.
PR 09-OCT-2001; 2001US-0328044P.
PR 09-OCT-2001; 2001US-0328056P.
PR 12-OCT-2001; 2001US-0328849P.
PR 15-OCT-2001; 2001US-0329414P.
PR 17-OCT-2001; 2001US-0330142P.
PR 22-OCT-2001; 2001US-0341058P.
PR 24-OCT-2001; 2001US-0339266P.
PR 24-OCT-2001; 2001US-0343629P.
PR 25-OCT-2001; 2001US-0349575P.
PR 01-NOV-2001; 2001US-0346357P.
PR 12-APR-2002; 2002US-0371972P.
PR 13-APR-2002; 2002US-0371980P.
PR 19-APR-2002; 2002US-0373261P.
PR 19-APR-2002; 2002US-0373805P.
PR 23-APR-2002; 2002US-0374738P.
PR 16-MAY-2002; 2002US-0381101P.
PR 17-MAY-2002; 2002US-0381635P.
PR 29-MAY-2002; 2002US-0383830P.
PR 01-OCT-2002; 2002US-00262839.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Alsobrook JP, Anderson DW, Boidog FL, Burgess CE, Catterton E,
PI Edinger SR, Ellerman K, Gerlach VL, Gorman L, Guo X, Ji W,
PI Kekuda R, Leach MD, Li L, Miller CE, Patturajan M, Rieger DK,
PI Rothenberg ME, Shinkets RA, Smithson G, Spytek KA, Taupier RJ,
PI Vernet CAM, Voss EZ, Zerhusen BD, Zhong M,
XX
XX WPI; 2003-381625/36.
DR N-PSDB; ACC72077.
XX
XX NOVX polypeptides and nucleic acids useful for diagnosing, preventing or
PT treating NOVX-associated disorders, e.g. diabetes, obesity, cancer or
PT dyslipidemia, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
XX Claim 1; Page 107; 487pp; English.
XX
XX The present invention relates to novel human NOV proteins and their
CC coding sequences (ACC72075-ACC72181 and ABR58363-ABR58469). The NOV
CC proteins are useful in manufacturing a medicament for treating a syndrome
CC associated with a human disease. The NOV proteins and coding sequences
CC may be used to diagnose, treat or prevent metabolic disorders such as
CC diabetes or obesity, infections, cachexia, cancer, neurodegenerative
CC disorders such as Alzheimer's disease or Parkinson's disease, immune
CC disorders, haematopoietic disorders and various dyslipidaemias
XX
XX Sequence 1120 AA;
Query Match 63.4%; Score 753; DB 6; Length 1120;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 159 CQTYMDIVIVLDGNSIPYVEVQHFILNLIKFFYIGPGQIQGVQYGEDVVEHFLND 218
DB 90 CQTYMDIVIVLDGNSIPYVEVQHFILNLIKFFYIGPGQIQGVQYGEDVVEHFLND 149
QY 219 YRSVKDVVEAASHTEQRGGTETRTAFGIEPARSEAFQGGKKGAKKVMIVITDGHSDSP 278
DB 150 YRSVKDVVEAASHTEQRGGTETRTAFGIEPARSEAFQGGKKGAKKVMIVITDGHSDSP 209
QY 279 DLEKVIQOESRDNTTRAVAVLGYNNRGINPETFLEIKYIASDPDDKHFNVTDEAAL 338
DB 210 DLEKVIQOESRDNTTRAVAVLGYNNRGINPETFLEIKYIASDPDDKHFNVTDEAAL 269

QY 339 KDIVDALGDRIFSLGKNETSFGLMSQTSFGSSHVVDGVLGAVGAYDWMGAVLKET 398
DB 270 KDIVDALGDRIFSLGKNETSFGLMSQTSFGSSHVVDGVLGAVGAYDWMGAVLKET 329
QY 399 SAGKVIPLRESYLKEFPEELKNHGAYLGYVTVTSVSSRQGRVYVAGAPRNFHTGKVLFT 458
DB 330 SAGKVIPLRESYLKEFPEELKNHGAYLGYVTVTSVSSRQGRVYVAGAPRNFHTGKVLFT 389
QY 459 MHNRSLLTIHOAMEGQOIGSYFGSEITSVDIDGGVTDVLLVGAPMYFNGREGKQVYV 518
DB 390 MHNRSLLTIHOAMEGQOIGSYFGSEITSVDIDGGVTDVLLVGAPMYFNGREGKQVYV 449
QY 519 ELRQNRFFVYNGTLKDSHSYQNAFEGSSIASVRDLNQSNDVNVVVGAPLEDNHAGAIYIFH 578
DB 450 ELRQNRFFVYNGTLKDSHSYQNAFEGSSIASVRDLNQSNDVNVVVGAPLEDNHAGAIYIFH 509
QY 579 GFRGSIILKTPKQRTITASELATGLOFQCSIHGQIDLNEDGLIDLAVGALGNVILWSPV 638
DB 510 GFRGSIILKTPKQRTITASELATGLOFQCSIHGQIDLNEDGLIDLAVGALGNVILWSPV 569
QY 639 VQINASHFEPKINIFHRDCKRSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMD 698
DB 570 VQINASHFEPKINIFHRDCKRSGRDATCLAAFLCFTPIFLAPHFQTTTIGIRYNATMD 629
QY 699 ERYTPRAHLDEGGDRFTNRAVLLSSGOELCERINFHVLTADYVXKVPFVSVEYSLDDPD 758
DB 630 ERYTPRAHLDEGGDRFTNRAVLLSSGOELCERINFHVLTADYVXKVPFVSVEYSLDDPD 689
QY 759 HGPMLDDGWPPTTLRVSVFPFNGCNEDEHCVDPDLVDARSDLPTAMEYQCVLRKPAQDCS 818
DB 690 HGPMLDDGWPPTTLRVSVFPFNGCNEDEHCVDPDLVDARSDLPTAMEYQCVLRKPAQDCS 749
QY 819 AYTLSFDTTTFIESTQRVAEATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGS 878
DB 750 AYTLSFDTTTFIESTQRVAEATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGS 809
QY 879 IECVNEERRLOKQVNVSYPPFFRAKAKVAFRLD 911
DB 810 IECVNEERRLOKQVNVSYPPFFRAKAKVAFRLD 842
RESULT 17
AAU19663
ID AAU19663 standard; protein; 707 AA.
XX AC AAU19663;
XX DT 04-DEC-2001 (first entry)
XX DE Human novel extracellular matrix protein, Seq ID No 313.
XX KW Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;
XX antianemic; antirheumatic; antisclerotic; cardiant; vascular;
XX cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
XX antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
XX human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
XX cancers; hyperproliferative disorder; breast neoplasm; melanoma;
XX Sezary syndrome; Gaucher's disease; neurological diseases;
XX Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
XX cardiac arrest; tachycardia; angina; infection; corneal infections;
XX wound healing; immunogen; gene therapy; antisenase; food additive.
XX OS Homo sapiens.
XX PN WO200155368-A1.
XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001348.
XX PR 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209457P.
 PR 28-JUN-2000; 2000US-0214896P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216680P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226868P.
 PR 23-AUG-2000; 2000US-0227182P.
 PR 30-AUG-2000; 2000US-0227009P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231143P.
 PR 08-SEP-2000; 2000US-0231141P.
 PR 08-SEP-2000; 2000US-0231080P.
 PR 08-SEP-2000; 2000US-0231081P.
 PR 14-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239935P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0241826P.
 PR 08-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 08-NOV-2000; 2000US-0246613P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249219P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249246P.
 PR 17-NOV-2000; 2000US-0249255P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 01-DEC-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251866P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 11-DEC-2000; 2000US-0251990P.
 PR 05-JAN-2001; 2001US-0259678P.
 (HUMA-) HUMAN GENOME SCI INC.
 Rosen CA, Barash SC, Ruben SM;
 WPI; 2001-465572/50.
 N-PSDB; AAS31234.
 Nucleic acid molecules encoding human secreted extracellular matrix proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers.

PS Claim 11; SEQ ID NO 313; 577bp; English.
XX The invention relates to isolated nucleic acid molecules encoding novel
CC human secreted extracellular matrix proteins (SPs). The polynucleotides
CC and proteins are used to prevent, treat a medical condition in e.g.
CC humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For
CC example, disorders associated with decreased expression of SPs. The SP
CC polynucleotide or a vector expressing them may be administered to treat
CC diseases by gene therapy. Antisense molecules may be administered to down
CC regulate expression of SPs by binding with the cells own genes and
CC preventing their expression. The polynucleotides may also be used as DNA
CC probes in diagnostic assays. The SPs may also be used as antigens to
CC produce antibodies and to identify modulators (agonists and antagonists)
CC of the SPs. The anti-(sp) antibodies and antagonists may also be used to
CC down regulate expression and activity of SP and as diagnostic agents for
CC detecting the presence of SPs in samples. The disorders include for
CC example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency
CC virus) infections, anemia, rheumatoid arthritis and multiple sclerosis),
CC cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of
CC the breast or liver, Sezary syndrome and Gaucher's disease), neurological
CC diseases (e.g. Alzheimer's disease, Parkinson's disease) cardio-
CC /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),
CC infections caused by bacteria, viruses and fungi and ocular disorders
CC (e.g. corneal infections). Other uses include wound healing, maintenance
CC of organs before transplantation, support of cell culture of primary

Query Match 41.2%; Score 489; DB 4; Length 707;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
Qy 1 MDLPRLGVAVWALSINPFGFTDTNMDTRKPRVIGSRTAFFGYTVQOHDISGNKMLVGA 60
Db 17 MDLPRLGVAVWALSINPFGFTDTNMDTRKPRVIGSRTAFFGYTVQOHDISGNKMLVGA 76
Qy 61 PLETNGYQKTDVYKCPVHGNCTKLNGLRVTLNSVSEKDNKRLGLSLATPNKNSFLA 120
Db 77 PLETNGYQKTDVYKCPVHGNCTKLNGLRVTLNSVSEKDNKRLGLSLATPNKNSFLA 136
Qy 121 CSPLWSHECGSSYYTTCMCSSRVNSFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db 137 CSPLWSHECGSSYYTTCMCSSRVNSFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 196
Qy 181 VQFLINILKFFIGPQIQGVVQYGEDVVFHFLNDYRSVDVVAASHIEQRGTTET 240
Db 197 VQFLINILKFFIGPQIQGVVQYGEDVVFHFLNDYRSVDVVAASHIEQRGTTET 256
Qy 241 RTAFGIEFARSEAFQKGRGAKKVMIVITDGHSDSPDLKVIQOSRDNVTRYAVAVL 300
Db 257 RTAFGIEFARSEAFQKGRGAKKVMIVITDGHSDSPDLKVIQOSRDNVTRYAVAVL 316
Qy 301 GYNNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGTTKNET 360
Db 317 GYNNRRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDIDVALGDRIFSLGTTKNET 376
Qy 361 SFGLEMSQTFSSHVVEDGVLGAVGAYDNKAVLKETSAGKVIPLRESYLKEFFPEELKN 420
Db 377 SFGLEMSQTFSSHVVEDGVLGAVGAYDNKAVLKETSAGKVIPLRESYLKEFFPEELKN 436
Qy 421 HGAYLGYTVTSVSSRGRVYVAGAPRNFHTGKVLFTMNNRSLTIHQAMRQQQIGSYF 480
Db 437 HGAYLGYTVTSVSSRGRVYVAGAPRNFHTGKVLFTMNNRSLTIHQAMRQQQIGSYF 496
Qy 481 GSEITSDVIDDGDVTDVLLGAPMYFNEGRGKVVYELRQNFVYNGTLKDSHSYQNA 540
Db 497 GSEITSDVIDDGDVTDVLLGAPMYFNEGRGKVVYELRQNFVYNGTLKDSHSYQNA 556
Qy 541 RFGSSIASVRDLNODSVNDVVGAPLEDNHAGAIYIFHGRGSLTKPKQRIITASELATG 600
Db 557 RFGSSIASVRDLNODSVNDVVGAPLEDNHAGAIYIFHGRGSLTKPKQRIITASELATG 616
Qy 601 LQYFGCSIHGOLDINEGLDIDLVAGLGNVILWSRPVQINASLHPEPSKINIIFRDCK 660
Db 617 LQYFGCSIHGOLDINEGLDIDLVAGLGNVILWSRPVQINASLHPEPSKINIIFRDCK 676

Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 691
Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 707
RESULT 18
ABP47883
ID ABP47883 standard; protein; 707 AA.
AC ABP47883;
DT 23-AUG-2002 (first entry)
DE Human polypeptide SEQ ID NO 313.
KW Human; nontropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antinflammatory; anti-HIV; antibacterial; vulnery;
KW antiparkinsonian; antitickling; antianemic; antiarthritic; cancer;
KW antihemetic; hepatocytic; cerebroprotective; antinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiac; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX Homo sapiens.
XX US2002042386-A1.
XX 11-APR-2002.
XX 17-JAN-2001; 2001US-00764870.
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0220963P.
PR 16-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229511P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.
PR 20-OCT-2000; 2000US-0240950P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
PA (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2002-470713/50.
XX N-PSDB; ABQ66558.
XX
XX New nucleic acid encoding human proteins, useful for diagnosis, treatment
XX and prevention of e.g. osteoporosis, also related polypeptides and
XX antibodies.
XX
XX Claim 11; SEQ ID NO 313; 235pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABQ66521-ABQ66785) and proteins
XX (ABP47846-ABP48110) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer, and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from USPTO at seqdata.uspto.gov/sequence.html?DocID=999909764870
XX
XX Sequence 707 AA;
XX
Query Match 41.2%; Score 489; DB 5; Length 707;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDLPRLVAVAWLSLWPGFTDTFNMDTRKPRVTPGSRTPAFGVTYQOCHDSGNKWLWVGA 60
DB 17 MDLPRLVAVAWLSLWPGFTDTFNMDTRKPRVTPGSRTPAFGVTYQOCHDSGNKWLWVGA 76
QY 61 PLETNGYQKTDGYKCPVHNGCTKLNLRVLSNVSRKDNRLGLSLATNPKNDSFLA 120
DB 77 PLETNGYQKTDGYKCPVHNGCTKLNLRVLSNVSRKDNRLGLSLATNPKNDSFLA 136
QY 121 CSPLWSEHCSSYTTGMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 180
DB 137 CSPLWSEHCSSYTTGMSRVNSNFRSKTVAPALQRCQTYMDIVIVLDGNSNLYPWVE 196
QY 181 VQHFLINLKFFYIGPQIQGVVQYGEDVHEFHNLNDRSVKDVVEAASHIEQGGTET 240
DB 197 VQHFLINLKFFYIGPQIQGVVQYGEDVHEFHNLNDRSVKDVVEAASHIEQGGTET 256
QY 241 RTAFGLFEARSEAFQGGKGAKKMIVITDGHSDSPLEKVIQOQSRDNTVYAVAVL 300
DB 257 RTAFGLFEARSEAFQGGKGAKKMIVITDGHSDSPLEKVIQOQSRDNTVYAVAVL 316
QY 301 GYNNRRGINPETFLNEIKYIASDPDDKHFNFVNTDEAALKDIDVALGDRIFSLGKNET 360
DB 317 GYNNRRGINPETFLNEIKYIASDPDDKHFNFVNTDEAALKDIDVALGDRIFSLGKNET 376

QY 361 SFGLEMSOTGFSSHVEDGVLLGAVGAYDMNGAVLKTSAGKVIPLRESYLKEPPELKN 420
DB 377 SFGLEMSOTGFSSHVEDGVLLGAVGAYDMNGAVLKTSAGKVIPLRESYLKEPPELKN 436
QY 421 HGAYLGYTVTSVSSRQGRVVVAGAPRNFHTKGVILFTMNNRSLTIHQAVRQGOIGSYF 480
DB 437 HGAYLGYTVTSVSSRQGRVVVAGAPRNFHTKGVILFTMNNRSLTIHQAVRQGOIGSYF 496
QY 481 GSEITSVDIDGVDYDVLVLCAPMYFNEGRGKVVYVELRQNFVYNGTLKDSHSYQNA 540
DB 497 GSEITSVDIDGVDYDVLVLCAPMYFNEGRGKVVYVELRQNFVYNGTLKDSHSYQNA 556
QY 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIHFGRFGSILKTPKQRTITASELATG 600
DB 557 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIHFGRFGSILKTPKQRTITASELATG 616
QY 601 LQYFGCSIHGOLDLNEEDGLDLAVGALGNVILMSRPVQINASLHPEPSKINIFHRDCK 660
DB 617 LQYFGCSIHGOLDLNEEDGLDLAVGALGNVILMSRPVQINASLHPEPSKINIFHRDCK 676
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 691
DB 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGGI 707
RESULT 19
ADCL0845
ID ADCL0845 standard; protein; 707 AA.
XX
XX AC ADCL0845;
XX
XX DT 18-DEC-2003 (first entry)
XX Human extracellular matrix protein from gene 38.
XX Extracellular matrix protein; cytostatic; antibacterial; virucide;
XX neuroprotective; gynaecological; gastrointestinal-Gen; caidiand;
XX cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
XX respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
XX neotropic; anti-allergic; cancer; bacterial infection; viral infection;
XX neural disorder; immune system disorder; blood disorder;
XX muscular disorder; reproductive disorder; gastrointestinal disorder;
XX pulmonary disorder; cardiovascular disorder; renal disorder;
XX inflammatory disorder; proliferative disorder; Human.
XX
XX Homo sapiens.
XX
XX US2003059875-A1.
XX
XX 27-MAR-2003.
XX
XX 19-APR-2002; 2002US-00125540.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0204967P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216847P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.

CC protein, the gene corresponding to the cDNA sequence, and identifying an
 CC activity in a biological assay (comprising expressing the nucleic acid in
 CC a cell, isolating the supernatant, detecting an activity in a biological
 CC assay and identifying the protein in the supernatant having the
 CC activity). The nucleic acids and proteins display the following
 CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective,
 CC Gynaecological, Gastrointestinal, Cardiant, Cardiovascular-Gen,
 CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 41.2%; Score 489; DB 7; Length 707;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 MDLPRGLVWAWALSLWPGFTDTFNMTRKPRVTPGRTAFPGYTVQOHDISGNKWLWGA 60
 Db 17 MDLPRGLVWAWALSLWPGFTDTFNMTRKPRVTPGRTAFPGYTVQOHDISGNKWLWGA 76
 Qy 61 PLENTGQKTDGVVYKCPVHGNCTKLNGLRVTLNVSERKDNMRLGLSLATPKNSFLA 120
 Db 77 PLENTGQKTDGVVYKCPVHGNCTKLNGLRVTLNVSERKDNMRLGLSLATPKNSFLA 136
 Qy 121 CSPLWSHECGSSYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
 Db 137 CSPLWSHECGSSYTTGMCGRVNSNFRFSKTVAPALQRCQTYMDIVIVLDGNSIYPWVE 196
 Qy 181 VQHFNLINLKFFIGPQIOGVVQVGEDVVFHFHNDYRSVVDVVEAASHIEQRGSTT 240
 Db 197 VQHFNLINLKFFIGPQIOGVVQVGEDVVFHFHNDYRSVVDVVEAASHIEQRGSTT 256
 Qy 241 RTAFGIEFARSEAFQKGRGKAKMWLITDGHSDSPLEKVIQOSRDNTRVAVL 300
 Db 257 RTAFGIEFARSEAFQKGRGKAKMWLITDGHSDSPLEKVIQOSRDNTRVAVL 316
 Qy 301 GYNRRGINPETFLNEIKYIASDPDDKHFNFVTDAAALKDIDALGDRIFSLEGTKNET 360
 Db 317 GYNRRGINPETFLNEIKYIASDPDDKHFNFVTDAAALKDIDALGDRIFSLEGTKNET 376
 Qy 361 SFGLEMSQTFSSHVEDGVLLGAVAYDNGAVLAKETSAKVIPLRESYLKEFFPEELKN 420
 Db 377 SFGLEMSQTFSSHVEDGVLLGAVAYDNGAVLAKETSAKVIPLRESYLKEFFPEELKN 436
 Qy 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNHTGKVLFTMNNRSLTIHOAMRGOQIGSYF 480
 Db 437 HGAYLGYTVTSVSSRQGRVYVAGAPRNHTGKVLFTMNNRSLTIHOAMRGOQIGSYF 496
 Qy 481 GSITSDVDIDGCVTDVLLGAPWYNEGREKGVVYELRONRFVYNGTLKDSHYQNA 540
 Db 497 GSITSDVDIDGCVTDVLLGAPWYNEGREKGVVYELRONRFVYNGTLKDSHYQNA 556
 Qy 541 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFGRSILKTPKORITASELATG 600
 Db 557 RFGSSIASVRDLNQDSYNDVVGAPLEDNHAGAIYIFHGFGRSILKTPKORITASELATG 616
 Qy 601 LQYFGCSIHQDLNEDGLDIAVAGALGNVILWSPVQINASLHFFPSKINIFHRDCK 660
 Db 617 LQYFGCSIHQDLNEDGLDIAVAGALGNVILWSPVQINASLHFFPSKINIFHRDCK 676
 Qy 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 691
 Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 707

RESULT 20

AAU76854
ID AAU76854 standard; protein; 193 AA.

XX AAU76854;

XX 21-MAY-2002 (first entry)

XX Human integrin alpha subunit Alpha 11 A domain.

XX Human; integrin alpha subunit; A domain; Alpha 11; integrin beta subunit;

KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
 KW ischaemia-reperfusion injury; immune complex; parasitic disease;
 KW antiinflammatory; vasotropic; antiparasitic; vulnary; gene therapy.

XX Homo sapiens.

XX WO200209737-A1.

XX 07-FEB-2002.

XX 31-JUL-2001; 2001WO-US023957.

XX 31-JUL-2000; 2000US-0221950P.

XX 11-JAN-2001; 2001US-00758493.

XX 13-MAR-2001; 2001US-00805354.

XX (GSHO) GEN HOSPITAL CORP.

XX Arnaout AM, Li R, Xiong J;

XX WPI; 2002-188687/24.

Novel high affinity integrin polypeptide useful for treating restenosis
 and parasitic diseases, comprises all or part of variant integrin alpha
 subunit A domain or variant integrin beta subunit A-like domain.

Example 2; Fig 5; 55pp; English.

The invention relates to a high affinity integrin polypeptide comprising
 all or part of a variant integrin alpha subunit A domain or a variant
 integrin beta subunit A-like domain. The polypeptide, preferably the
 CD11b alpha subunit A domain, where I at residue 332 has been replaced by
 G or A, F at residue 313 and A at residue 320 have been replaced by C, or
 V at residue 315 and A at residue 320 have been replaced by C, is useful
 for determining if a test compound is a candidate compound for binding to
 CD11b or for treating an inflammatory disorder, by contacting a test
 compound with the polypeptide and determining if the test compound binds
 to the polypeptide. The integrin subunits are useful for reducing
 skeletal muscle injury, for treating disorders caused by ischaemia-
 reperfusion injury, immune complexes, restenosis and parasitic diseases,
 to purify variant integrin polypeptide ligands and as bait proteins in
 two-hybrid or three-hybrid assays. This sequence represents the human
 integrin alpha subunit Alpha 11 A domain

Sequence 193 AA;

Query Match 16.2%; Score 193; DB 5; Length 193;
 Best Local Similarity 100.0%; Pred. No. 2.1e-182;

Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQTMDIVIVLDGNSIYPWVEVQHFLINILKXFYIGPQIOGVVQVGEDVVFHFLND 218

Db 1 CQTMDIVIVLDGNSIYPWVEVQHFLINILKXFYIGPQIOGVVQVGEDVVFHFLND 60

Qy 219 YRSVKDVVEAASHIEQRGSTTETRTAFGIEFARSEAFQKGRGKAKMWITDGHSDSP 278

Db 61 YRSVKDVVEAASHIEQRGSTTETRTAFGIEFARSEAFQKGRGKAKMWITDGHSDSP 120

Qy 279 DLEKVIQOSRDNTRVAVAVLGYNNRGINPETFLNEIKYIASDDDDKGFNFVTDAAAL 338

Db 121 DLEKVIQOSRDNTRVAVAVLGYNNRGINPETFLNEIKYIASDDDDKGFNFVTDAAAL 180

Qy 339 KDIVDALGDRIFS 351

Db 181 KDIVDALGDRIFS 193

RESULT 21

AAU76863

ID AAU76863 standard; protein; 193 AA.

XX AAU76863;

DT 21-MAY-2002 (first entry)
DE Human integrin alpha subunit Alpha 11 variant A domain.
XX
XX Human; integrin alpha subunit; A domain; Alpha 11; integrin beta subunit;
KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;
KW antiinflammatory; vasotropic; antiparasitic; vulnery; gene therapy;
KW mutin.
XX
XX Homo sapiens.
OS Synthetic.
OS
XX
XX
FH Key Location/Qualifiers
FT Misc-difference 191
FT /note= "Wild-type Ile substituted by any other amino
FT acid"
XX
XX WO200209737-A1.
XX
XX 07-FEB-2002.
XX
XX 31-JUL-2001; 2001WO-US023957.
XX
XX 31-JUL-2000; 2000US-0221950P.
XX 11-JAN-2001; 2001US-00758493.
XX 13-MAR-2001; 2001US-00805354.
XX
XX (GEO) GEN HOSPITAL CORP.
XX
XX Arnaout AM, Li R, Xiong J;
XX
XX WPI; 2002-188687/24.
XX
XX Novel high affinity integrin polypeptide useful for treating restenosis
PT and parasitic diseases, comprises all or part of variant integrin alpha
PT subunit A domain or variant integrin beta subunit A-like domain.
XX
XX Claim 57; Page; 55pp; English.
XX
XX The invention relates to a high affinity integrin polypeptide comprising
CC all or part of a variant integrin alpha subunit A domain or a variant
CC integrin beta subunit A-like domain. The polypeptide, preferably the
CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
CC V at residue 315 and A at residue 320 have been replaced by C, is useful
CC for determining if a test compound is a candidate compound for binding to
CC CD11b or for treating an inflammatory disorder, by contacting a test
CC compound with the polypeptide and determining if the test compound binds
CC to the polypeptide. The integrin subunits are useful for reducing
CC skeletal muscle injury, for treating disorders caused by ischaemia-
CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
CC to purify variant integrin polypeptide ligands and as bait proteins in
CC two-hybrid or three-hybrid assays. This sequence represents a human
CC integrin alpha subunit Alpha 11 variant A domain. Note: This variant
CC sequence is not featured in the specification but has been derived from
CC the wild-type protein shown in AAU76854
XX
SQ Sequence 193 AA;
Query Match 16.0%; Score 190; DB 5; Length 193;
Best Local Similarity 100.0%; Pred. No. 2e-179;
Matches 190; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 159 CQYMDIVVLVLDGNSIYPVWEVQHFLNLIKFIQGGQIQGVVQYGEDVVEHFLND 218
Dy 1 CQYMDIVVLVLDGNSIYPVWEVQHFLNLIKFIQGGQIQGVVQYGEDVVEHFLND 60
Qy 219 YRSKDVVEASHIEQKGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGHSDSP 278
Dy 61 YRSKDVVEASHIEQKGGTETRTAFGIEFARSEAFQGGKGAKKVMIVITDGHSDSP 120
Qy 279 DLEKVIQOSERDNTVRVAVAVLGYNNRRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 338

Db 121 DLEKVIQOSERDNTVRVAVAVLGYNNRRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 180
Qy 339 KDIVDALGDR 348
Dy 181 KDIVDALGDR 190
RESULT 22
AAB50087
ID AAB50087 standard; protein; 1188 AA.
XX
XX AAB50087;
XX
XX 19-MAR-2001 (first entry)
XX Murine A259.
XX Human; A259; integrin alpha subunit; fibrosis; liver disease; lung;
KW kidney; cancer; osteoporosis; acute myeloid leukaemia; HIV infection;
KW rheumatoid arthritis.
XX
XX Mus sp.
XX
XX OS
XX FH Key Location/Qualifiers
FT Domain 1..1141
FT /label= Extracellular_domain
FT Peptide 1..22
FT /label= Signal_peptide
FT Protein 23..1188
FT /label= Mature_protein
FT Domain 39..74
FT /label= Integrin_alphasubunit_repeat_domain_#1
FT Domain 115..157
FT /label= Integrin_alphasubunit_repeat_domain_#2
FT Domain 164..345
FT /label= I_domain
FT Domain 367..392
FT /label= Integrin_alphasubunit_repeat_domain_#3
FT Domain 421..455
FT /label= Integrin_alphasubunit_repeat_domain_#4
FT Domain 478..516
FT /label= Integrin_alphasubunit_repeat_domain_#5
FT Domain 540..575
FT /label= Integrin_alphasubunit_repeat_domain_#6
FT Domain 602..640
FT /label= Integrin_alphasubunit_repeat_domain_#7
FT Domain 1142..1164
FT /label= Transmembrane_domain
FT Domain 1165..1188
FT /label= Cytoplasmic_domain
XX
XX WO200073339-A1.
XX
XX 07-DEC-2000.
XX
XX 15-MAY-2000; 2000WO-US013262.
XX
XX 28-MAY-1999; 99US-00322790.
XX 27-APR-2000; 2000US-00561263.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX Pan Y, Lora JM;
XX
XX WPI; 2001-041142/05.
XX N-PSDB; AAC91904, AAC91905.
XX
XX Nucleic acid encoding alpha-integrin subunits, useful for treatment and
PT diagnosis of fibrosis, e.g. of the liver.
XX
XX Claim 8; Fig 5; 164pp; English.
XX

CC The present sequence is murine integrin alpha subunit, A259. A259 is
 CC homologous with the alpha1 and alpha10 integrin subunits and is
 CC overexpressed in fibrosis. A259 is implicated in regulation of
 CC proliferation, differentiation and/or function of many different cell
 CC types. Inhibitors of A259 activity are useful for the treatment of liver
 CC disease, particularly fibrosis, and also fibrosis in other organs
 CC (specifically lung and kidney). In addition, A259 can be used for
 CC treatment and prevention of cancer, osteoporosis, acute myeloid
 CC leukaemia, HIV infection, and rheumatoid arthritis
 XX
 XX SQ Sequence 1188 AA;

Query Match 10.3%; Score 122; DB 4; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 2.9e-111; Indels 0; Gaps 0;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GNCCTKLNGLGRVTLNVSERKDNRLGLSLATNPKDSFLACSPLWSHECGSSYYTTGMC 140
 DB 81 GNCCTKLNGLGRVTLNVSERKDNRLGLSLATNPKDSFLACSPLWSHECGSSYYTTGMC 140

QY 141 RVNSNFRFSKTVPALQRCQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGGQIQ 200
 DB 141 RVNSNFRFSKTVPALQRCQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGGQIQ 200

QY 201 VG 202
 DB 201 VG 202

RESULT 23
 AAU10552
 ID AAU10552 standard; protein; 1188 AA.
 XX
 AC AAU10552;
 XX
 DT 14-FEB-2002 (first entry)
 XX
 DE Murine A259 polypeptide.

XX Mouse; A259; integrin alpha subunit; integrin alpha 10; secreted protein;
 XX liver disease; fibrosis; lung; kidney; bone associated disorder; blood;
 KW cartilage associated disorder; haematopoietic disorder; bone marrow;
 KW immune related disease; apoptotic disorder; neuronal tissue disease;
 KW neurodegenerative disease; gene therapy; cancer; cytostatic; osteopathic;
 KW nephrotropic; immunomodulator; anti-inflammatory; neuroprotective;
 KW antiarthritic; antianemic; antiallergic; antiasthmatic; dermatological;
 KW antidiabetic; anticonvulsant; antiparkinsonian.
 XX
 OS Mus musculus.

XX FH Key Location/Qualifiers
 FT Domain 1. .1141
 FT Peptide /note= "Extracellular domain"
 FT Protein /note= "Signal peptide"
 FT Domain 23. .1188
 FT Domain /note= "Mature murine A259"
 FT Domain 39. .74
 FT Domain /note= "Integrin alpha repeat domain"
 FT Domain 115. .157
 FT Domain /note= "Integrin alpha repeat domain"
 FT Domain 164. .345
 FT Domain /note= "Integrin alpha repeat domain"
 FT Domain 367. .392
 FT Domain /note= "I domain or Von Willebrand Factor type A domain"
 FT Domain 421. .455
 FT Domain /note= "Integrin alpha repeat domain"
 FT Domain 478. .516
 FT Domain /note= "Integrin alpha repeat domain"
 FT Domain 540. .575
 FT Domain /note= "Integrin alpha repeat domain"
 FT Domain 602. .640
 FT Domain /note= "Integrin alpha repeat domain"

FT Domain 1142. .1164
 FT /note= "Transmembrane domain"
 FT Domain 1165. .1188
 FT /note= "Cytoplasmic domain"
 XX WO200181414-A2.
 XX 01-NOV-2001.
 XX 27-APR-2001; 2001WO-US013516.
 XX 27-APR-2000; 2000US-00561263.
 XX (MILL-) MILLENNIUM PHARM INC.
 XX Pan Y, Lora J;
 DR WPI; 2002-041397/05.
 DR N-PSDB; AAS16874.
 XX New A259 nucleic acids and polypeptides, which comprise integrin alpha
 FT subunit, useful for diagnosing, preventing or treating e.g. liver
 FT disease, kidney or lung fibrosis, cancers, blood disorders or immune
 FT related diseases.
 XX
 PS Claim 9; Fig 5; 168pp; English.
 XX
 CC The invention relates to human and murine A259 nucleic acid molecules
 CC which encode secreted proteins with homology to integrin alpha subunits,
 CC specifically to integrin alpha 10. The A259 polypeptide and nucleic acid
 CC are useful for treating liver disease or fibrosis, particularly kidney
 CC fibrosis or lung fibrosis. The A259 polypeptide and nucleic acid are also
 CC useful for diagnosing, preventing or treating cartilage and bone
 CC associated disorders (such as bone cancer, achondroplasia, myeloma,
 CC fibrous dysplasia, scoliosis, osteoarthritis, osteosarcoma and
 CC osteoporosis), bone marrow, blood and haematopoietic disorders (such as
 CC acute myeloid leukaemia, haemophilia, anaemia and thalassaemia), immune
 CC related diseases (such as HIV, viral infections, cancers, T cell
 CC autoimmune disorders e.g. AIDS, and allergic inflammatory disorders e.g.
 CC asthma and psoriasis), apoptotic disorders (such as systemic lupus
 CC erythematosus and insulin-dependent diabetes mellitus), diseases of the
 CC neuronal tissues (such as epilepsy and muscular dystrophy) and
 CC neurodegenerative diseases (such as Parkinson's disease and Huntington's
 CC disease). This sequence represents the murine A259 polypeptide
 XX
 SQ Sequence 1188 AA;

Query Match 10.3%; Score 122; DB 5; Length 1188;
 Best Local Similarity 100.0%; Pred. No. 2.9e-111; Indels 0; Gaps 0;
 Matches 122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 81 GNCCTKLNGLGRVTLNVSERKDNRLGLSLATNPKDSFLACSPLWSHECGSSYYTTGMC 140
 DB 81 GNCCTKLNGLGRVTLNVSERKDNRLGLSLATNPKDSFLACSPLWSHECGSSYYTTGMC 140

QY 141 RVNSNFRFSKTVPALQRCQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGGQIQ 200
 DB 141 RVNSNFRFSKTVPALQRCQTYMDIVIVLDGNSIYPWVEVQHFLINILKFFYIGGQIQ 200

QY 201 VG 202
 DB 201 VG 202

RESULT 24
 ADE08585
 ID ADE08585 standard; protein; 360 AA.
 XX
 AC ADE08585;
 XX
 DT 29-JAN-2004 (first entry)
 XX
 DE Novel protein (useful for identifying genetic disorders) #740.

XX novel gene; novel protein; tissue marker; molecular weight marker;
KW chromosome marker; Genetic disorder.
XX Unidentified.
XX WO2003054152-A2.
XX 03-JUL-2003.
XX 10-DEC-2002; 2002WO-US039555.
XX 10-DEC-2001; 2001US-0339739P.
XX 11-DEC-2001; 2001US-0339453P.
XX 14-MAR-2002; 2002US-0365091P.
XX 14-MAR-2002; 2002US-0365384P.
XX 12-APR-2002; 2002US-0372381P.
XX 12-APR-2002; 2002US-0372615P.
XX 22-APR-2002; 2002US-00128558.
XX 24-APR-2002; 2002US-0376045P.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Asundi V, Goodrich RW, Ren F, Zhang J, Zhao QA, Wang J;
PI Ghosh M, Xue AJ, Wehrman T, Weng G, Zhou P, Drmanac RT, Wang Z;
PI Ma Y, Wang D, Chen R, Xu C, Boyle BJ;
XX WPI: 2003-569235/53.
XX N-PSDB; ADE07674.
XX New polynucleotides, useful for expressing recombinant proteins for
PT analysis, characterization or therapeutic use, or as markers for tissues
PT in which the corresponding protein is preferentially expressed.
XX Claim 20; SEQ ID NO 1651; 1177pp; English.
XX The invention comprises the amino acid and coding sequences of novel
CC proteins. The DNA and protein sequences of the invention are useful as:
CC markers for tissues in which the corresponding protein is preferentially
CC expressed; as molecular weight markers on gels; as chromosome markers or
CC tags; to identify chromosomes or to map related gene positions; and to
CC compare with endogenous DNA sequences in patients to identify potential
CC genetic disorders. The present amino acid sequence represents a protein
CC of the invention.
XX Sequence 360 AA;
SQ
Query Match 8.8%; Score 104; DB 7; Length 360;
Best Local Similarity 100.0%; Pred. No. 7.5e-94;
Matches 104; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 251 SEAFQKGRGKAKVMIVITDGHSHSPDLEKVIQOSERDNVTRYAVAVLGYNRRGINP 310
DB 232 SEAFQKGRGKAKVMIVITDGHSHSPDLEKVIQOSERDNVTRYAVAVLGYNRRGINP 291
QY 311 ETLFLNEIKYIASDPDDKHFFNVNTERALDKDIVDALGRIFSLRG 354
DB 292 ETLFLNEIKYIASDPDDKHFFNVNTERALDKDIVDALGRIFSLRG 335
RESULT 25
ABG66673
ID ABG66673 standard; protein; 109 AA.
XX AC ABG66673;
XX 30-AUG-2002 (first entry)
XX Human novel polypeptide #8.
XX Human, inflammatory condition; shock; sepsis; immune response; cancer;
KW wound healing; central nervous system disease; haematopoiesis;
KW peripheral nervous system disease; amyotrophic lateral sclerosis; tendon;

KW myeloid cell disorder; lymphoid cell disorder; platelet disorder; bone;
KW cartilage; ligament; nerve tissue; ulcer; osteoporosis; osteoarthritis;
KW bone degenerative disorder; periodontal disease; reperfusion injury;
KW lung fibrosis; liver fibrosis; autoimmune disorder; bacterial infection;
KW allergic condition; thrombolysis; thrombosis; coagulation disorder;
XX fungal infection.
XX Homo sapiens.
OS WO200244340-A2.
XX 06-JUN-2002.
XX 30-NOV-2001; 2001WO-US047004.
XX 30-NOV-2000; 2000US-00728952.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Wang J, Wang D;
PI Yamazaki V, Ujwal ML, Drmanac RT;
XX WPI: 2002-508509/54.
XX N-PSDB; ABK94897.
XX Novel nucleic acids and polypeptides for diagnosis, treatment of
PT inflammatory, autoimmune, nervous system, myeloid or lymphoid cell
PT disorders, cancer and promoting wound healing.
XX Claim 10; Page 572; 672pp; English.
XX The invention relates to human novel polynucleotides and associated
CC polypeptides. The polynucleotides and polypeptides are useful for
CC treating inflammatory conditions such as arthritis, nephritis, Crohn's
CC disease, ischaemia-reperfusion injury, shock, sepsis, immune responses
CC and cancer and for promoting wound healing. The sequences are used to
CC induce the proliferation of neural cells and regeneration of nerve and
CC brain tissue, and are useful for the treatment of central and peripheral
CC nervous system diseases and neuropathies, such as Alzheimer's disease,
CC Parkinson's disease, Huntington's disease and amyotrophic lateral
CC sclerosis. The sequences are involved in chemotactic or chemokinetic
CC activity, regulation of haematopoiesis, treatment of myeloid or lymphoid
CC cell disorders and platelet disorders such as thrombocytopenia,
CC regeneration of bone, cartilage, tendon, ligament and/or nerve tissue
CC growth, tissue repair, healing of burns, incisions, ulcers, treatment of
CC osteoporosis, osteoarthritis, bone degenerative disorders and periodontal
CC disease. The sequences of the invention are also useful for gut
CC protection or regeneration and treatment of lung or liver fibrosis,
CC reperfusion injury in various tissues, immune deficiencies and disorders
CC including severe combined immunodeficiency (SCID), bacterial or fungal
CC infections, autoimmune disorders e.g. multiple sclerosis and myasthenia
CC gravis, allergic conditions such as asthma, thrombolysis or thrombosis
CC and coagulation disorders. Sequences ABG66666-ABG66758 represent human
CC novel polypeptides of the invention
XX Sequence 109 AA;
SQ
Query Match 7.5%; Score 89; DB 5; Length 109;
Best Local Similarity 100.0%; Pred. No. 2e-79;
Matches 89; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MDLPRGLVVAVALSILWPGFTDTFNMDTRKPRVPGSRTAFFGYTVQQHDISGNKWLVVGA 60
DB 1 MDLPRGLVVAVALSILWPGFTDTFNMDTRKPRVPGSRTAFFGYTVQQHDISGNKWLVVGA 60
QY 61 PLETNGYQKTGDYVKCPVHGNCTKLNLG 89
DB 61 PLETNGYQKTGDYVKCPVHGNCTKLNLG 89
RESULT 26
AAU19822
ID AAU19822 standard; protein; 103 AA.

XX AC AAU19822;
XX DT 06-DEC-2001 (first entry)
XX DE Human novel extracellular matrix protein, Seq ID No 472.
XX KW Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;
KW anticancer; antirheumatic; antileukemic; antidiabetic; cardiac; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
KW anti-alzheimers; immune/autoimmune disease; HIV infection; anaemia;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disease; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infection; corneal infections;
KW wound healing; immunogen; gene therapy; antisense; food additive.
XX OS Homo sapiens.
XX PN WO200155368-A1.
XX FD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001348.
XX PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220363P.
PR 28-JUL-2000; 2000US-0220364P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228324P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235835P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.

17-NOV-2000; 2000US-02492997P.
17-NOV-2000; 2000US-0249299P.
17-NOV-2000; 2000US-0249300P.
01-DEC-2000; 2000US-0250160P.
01-DEC-2000; 2000US-0250391P.
05-DEC-2000; 2000US-0251030P.
05-DEC-2000; 2000US-0251988P.
05-DEC-2000; 2000US-0256719P.
06-DEC-2000; 2000US-0251479P.
08-DEC-2000; 2000US-0251856P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251869P.
08-DEC-2000; 2000US-0251989P.
11-DEC-2000; 2000US-0251990P.
11-DEC-2000; 2000US-0254097P.
05-JAN-2001; 2001US-0259678P.
(HUMA-) HUMAN GENOME SCI INC.
Rosen CA, Barash SC, Ruben SM;
WPI; 2001-465572/50.
N-PSDB; AAS31393.
Nucleic acid molecules encoding human secreted extracellular matrix proteins, used in preventing, treating or ameliorating a disorder, e.g. Alzheimer's and Parkinson's diseases and cancers.
Claim 11; SEQ ID NO 472; 577pp; English.
The invention relates to isolated nucleic acid molecules encoding novel human secreted extracellular matrix proteins (SPs). The polynucleotides and proteins are used to prevent, treat a medical condition in e.g. humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For example, disorders associated with decreased expression of SPs. The SP polynucleotide or a vector expressing them may be administered to treat diseases by gene therapy. Antisense molecules may be administered to treat regulate expression of SPs by binding with the cells own genes and preventing their expression. The polynucleotides may also be used as DNA probes in diagnostic assays. The SPs may also be used as antigens to produce antibodies and to identify modulators (agonists and antagonists) of the SPs. The anti-SP antibodies and antagonists may also be used to down regulate expression and activity of SP and as diagnostic agents for detecting the presence of SPs in samples. The disorders include for example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis), cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of the breast or liver, Sezary syndrome and Gaucher's disease), neurological diseases (e.g. Alzheimer's disease, Parkinson's disease) cardio-/cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina), infections caused by bacteria, viruses and fungi and ocular disorders (e.g. corneal infections). Other uses include wound healing, maintenance of organs before transplantation, support of cell culture of primary

Query Match 7.4%; Score 88; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 1.9e-78;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0

159 COTYMDIVILDGNSIYFWVEVGHFLINILKKFYIGFGQIQGVQYGEDVVHFEHLND 218
8 COTYMDIVILDGNSIYFWVEVGHFLINILKKFYIGFGQIQGVQYGEDVVHFEHLND 67

219 YFSVKDVEAASHIEQRGGTTRTAFGI 246
68 YRSVDVVEAASHIEQRGGTTRTAFGI 95

RESULT 27
AAU87675
ID AAU87675 standard; protein; 103 AA.
XX AC
XX AAU87675;
XX

```
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234597P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.

PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254037P.
PR 05-JAN-2001; 2001US-0259678P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-581633/65.
DR N-PSDB; ABK44005.
XX
PT New isolated nucleic acid encoding a protein for diagnosing, preventing,
PT treating or ameliorating medical conditions and used as food additives or
PT preservatives.
XX
PS Claim 9; SEQ ID NO 1193; 837pp; English.
XX
CC The invention describes an isolated nucleic acid molecule (I) encoding a
CC novel central nervous system protein. (I) and polypeptides (III) encoded
CC by (I), are used to treat a medical conditions and in diagnosis of a
CC pathological condition. Disorders which are diagnosed or treated include
CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease and
CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
CC adenocarcinomas and irritable bowel syndrome reproductive system,
CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
CC leukaemia, disorders involving neovascularisation e.g. malignancies,
CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
CC acute kidney failure and blood related disorders e.g. myocardial
CC infarction. The polypeptides can also be used to aid wound healing and
CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
CC maintain organs before transplantation, for supporting cell culture of
CC primary tissues, to regenerate tissues and in chemotaxis. The
CC polypeptides can also be used as a food additive or preservative to
CC increase or decrease storage capabilities, fat content, lipid, protein,
Query Match 7.4%; Score 88; DB 4; Length 103;
Best Local Similarity 100.0%; Pred No. 1.9e-78;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 159 CQTYMDIVIVLDGNSIYPMWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVVHPEFLND 218
Db 8 CQTYMDIVIVLDGNSIYPMWVEVQHFLINILKFKYIGPGQIQGVVQYGEDVVHPEFLND 67
Qy 219 YRSVKDVEEAASHIEQRGGTETRTAFGI 246
Db 68 YRSVKDVEEAASHIEQRGGTETRTAFGI 95
RESULT 28
ABP48042
ID ABP48042 standard; protein; 103 AA.
XX ABP48042;
XX 23-AUG-2002 (first entry)
XX Human polypeptide SEQ ID NO 472.
DE
```

XX Human; neutropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; anisickling; antianaemic; antiarthritic; cancer;
KW antirheumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antitumor; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
OS Homo sapiens.
XX
XX US2002042386-A1.
XX
XX 11-APR-2002.
XX
XX 17-JAN-2001; 2001US-00764870.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 28-JUN-2000; 2000US-0214886P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218202P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0228242P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239355P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;

XX WPI; 2002-470713/50.
DR N-PSDB; ABQ66717.
XX New nucleic acid encoding human proteins, useful for diagnosis, treatment
PT and prevention of e.g. osteoporosis, also related polypeptides and
PT antibodies.
XX Claim 11; SEQ ID NO 472; 235pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABQ66521-ABQ66785) and proteins
CC (ABP47846-ABP48110) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
CC infectious diseases such as viral, bacterial, fungal and parasitic
CC infections. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from USPTO at seqdata.uspto.gov/sequence.html?DocID=999909764870
XX
XX Sequence 103 AA;
XX
XX Query Match 7.4%; Score 88; DB 5; Length 103;
XX Best Local Similarity 100.0%; Pred. No. 1.9e-78;
XX Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 159 CQTYMDIVIVLDGNSIYPWVEVOHFLINILKFFYIGPGQIQGVQYGVGVHFLND 218
DB 8 CQTYMDIVIVLDGNSIYPWVEVOHFLINILKFFYIGPGQIQGVQYGVGVHFLND 67
QY 219 YRSVKDVVEAASHIEQRGGTETRTAFGI 246
DB 68 YRSVKDVVEAASHIEQRGGTETRTAFGI 95
RESULT 29
ADCI1004
ID ADCI1004 standard; protein; 103 AA.
XX
XX AC ADCI1004;
XX
XX 18-DEC-2003 (first entry)
XX
XX Human protein from extracellular matrix gene 38 #2.
XX
XX Extracellular matrix protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW neutropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; Human.
XX
XX Homo sapiens.
XX
XX US2003059875-A1.
XX
XX 27-MAR-2003.
XX
XX 19-APR-2002; 2002US-00125540.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217456P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226688P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228242P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234988P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249277P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 05-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 03-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764870.

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Ruben SM, Barash SC;

WPI; 2003-743765/70.

N-PSDB; ADC10739.

New isolated nucleic acids and polypeptides, useful for diagnosing, treating, and/or preventing disorders, such as cancer, infections, cardiovascular and inflammatory diseases.

KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
XX
XX Homo sapiens.
XX
XX WO200153312-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
XX
XX 21-JAN-2000; 2000US-00489725.
XX
XX 25-APR-2000; 2000US-0052317.
XX
XX 20-JUN-2000; 2000US-00598042.
XX
XX 19-JUL-2000; 2000US-00620312.
XX
XX 03-AUG-2000; 2000US-00653450.
XX
XX 14-SEP-2000; 2000US-00662191.
XX
XX 19-OCT-2000; 2000US-00693036.
XX
XX 23-NOV-2000; 2000US-00727344.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
XX Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
XX Zhou P, Goodrich R, Drmanac RT;
XX
XX WPI; 2001-442253/47.
XX
XX N-PSDB; AA159084.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
XX
XX Example 4; SEQ ID NO 3073; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AA157798-AA161369) and the
XX encoded polypeptides (AA38442-AA442213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: Immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
XX
XX Sequence 757 AA;
XX
Query Match 3.2%; Score 38; DB 4; Length 757;
Best Local Similarity 100.0%; Pred. No. 4.4e-28;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 738 DTADYKPVTFVSVEYSLEDDHGPMLDDGWPPTLRVSV 775
|||||
Db 76 DTADYKPVTFVSVEYSLEDDHGPMLDDGWPPTLRVSV 113
|||||
RESULT 35
ABG12950
ID ABG12950 standard; protein; 117 AA.
XX
XX AC ABG12950;
XX
XX DT 13-FEB-2002 (first entry)

XX Novel human diagnostic protein #12941.
DE
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
XX
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX 31-MAR-2000; 2000US-00540217.
XX
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
XX
XX N-PSDB; AA577137.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX
XX Claim 20; SEQ ID NO 43309; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
XX sequences. (I) is useful as hybridisation probes, polymerase chain
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
XX and in recombinant production of (II). The polynucleotides are also used
XX in diagnostics as expressed sequence tags for identifying expressed
XX genes. (I) is useful in gene therapy techniques to restore normal
XX activity of (II) or to treat disease states involving (II). (II) is
XX useful for generating antibodies against it, detecting or quantitating a
XX polypeptide in tissue, as molecular weight markers and as a food
XX supplement. (II) and its binding partners are useful in medical imaging
XX of sites expressing (II). (I) and (II) are useful for treating disorders
XX involving aberrant protein expression or biological activity. The
XX polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
XX amino acid sequences of the invention. Note: The sequence data for this
XX patent did not appear in the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 117 AA;
XX
Query Match 2.9%; Score 34; DB 4; Length 117;
Best Local Similarity 100.0%; Pred. No. 7.9e-25;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1130 FEISKQEDWQVPIWIVGSTIGLGLLLALLVLAL 1163
|||||
Db 42 FEISKQEDWQVPIWIVGSTIGLGLLLALLVLAL 75
|||||
RESULT 36
AAB25614
ID AAB25614 standard; peptide; 33 AA.
XX
XX AC AAB25614;
XX
XX DT 21-NOV-2000 (first entry)
XX
XX Human secreted protein ITGAl1 peptide encoded by secreted protein #6.

XX Secreted protein; immunosuppressant; anti-inflammatory; antiarthritic;
 KW antirheumatic; dermatological; antiproliferative; antiarteriosclerotic;
 KW anticancer; vulnery; antiviral; antibacterial; antifungal;
 KW immune disorder; Addison's disease; rheumatoid arthritis; dermatitis;
 KW multiple sclerosis; inflammatory disorder; inflammatory bowel disease;
 KW Crohn's disease; nephritis; hyperproliferative disorder;
 KW cardiovascular disorder; coronary arteriosclerosis; myocarditis; cancer;
 KW melanoma; lymphoma; wound healing; human.
 XX
 OS Homo sapiens.
 XX
 XX WO200029435-A1.
 XX
 XX 25-MAY-2000.
 XX
 XX 27-OCT-1999; 99WO-US025031.
 XX
 XX 28-OCT-1998; 98US-0105971P.
 XX
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX
 XX Ni J, Ruben SM, Olsen HS, Young PE, Kenny JJ, Moore PA, Wei Y;
 PI Greene JW;
 XX WPI; 2000-387742/33.
 XX
 XX Isolated nucleic acid molecules encoding human secreted proteins are used
 PT for the prevention, amelioration and treatment of autoimmune,
 PT inflammatory, hyperproliferative and cardiovascular disorders, cancer,
 PT wounds, and infectious diseases.
 XX
 XX Disclosure; Page 107; 803pp; English.
 XX
 XX The present invention relates to 12 secreted human proteins and the
 CC nucleotide sequences encoding them. The polynucleotide sequences given in
 CC AAA80606-A80623 encode the 12 secreted protein sequences given in
 CC AAB25576-B25593. The human secreted proteins have various activities
 CC dependent on the tissues in which they are expressed. Examples of the
 CC activities of the proteins include: immunosuppressant; anti-inflammatory;
 CC antirheumatic; antirheumatic; dermatological; antiproliferative;
 CC antiarteriosclerotic; anticancer; vulnery; antiviral; antibacterial;
 CC and antitumoral activity. The proteins, polypeptides, agonists and
 CC antagonists may be used to treat prevent and/or diagnose various disease,
 CC disorders and conditions examples of which include: immune disorders e.g.
 CC Addison's disease, rheumatoid arthritis, dermatitis, and multiple
 CC sclerosis; inflammatory disorders e.g. inflammatory bowel disease,
 CC Crohn's disease and nephritis; hyperproliferative disorders such as
 CC paraproteinemias and purpura; cardiovascular disorders e.g. coronary
 CC arteriosclerosis and myocarditis; cancer e.g. melanoma and lymphoma. The
 CC proteins and polynucleotide sequences may also be used in wound healing
 CC and the treatment of infectious diseases. The human secreted protein gene
 CC #6 and protein sequences are represented in sequences AAA80611 and
 CC AAB25581. Sequences AAA80650-A80651 represent genes related to the
 CC secreted protein gene#6
 XX
 XX Sequence 33 AA;
 SQ
 Query Match 2.0%; Score 33; DB 3; Length 33;
 Best Local Similarity 100.0%; Pred. No. 2.6e-24;
 Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 64 TNGYQKTDGVYKCPVTHGNTKLNLRVLSNV 96
 Db 1 TNGYQKTDGVYKCPVTHGNTKLNLRVLSNV 33
 RESULT 37
 ADA27114
 ID ADA27114 standard; peptide; 33 AA.
 XX
 AC ADA27114;
 XX

DT 20-NOV-2003 (first entry)
 XX Human novel secreted protein from gene 7 peptide.
 DE
 XX
 XX cytostatic; antiinflammatory; immunomodulator; neuroprotective;
 KW hemostatic; gene therapy; cancer; inflammation; immune disorder;
 KW neurological disorder; blood clotting disorder; food additive;
 KW preservative; human; secreted protein.
 XX
 OS Homo sapiens.
 XX
 XX US2003055231-A1.
 XX
 XX 20-MAR-2003.
 PD
 XX 29-OCT-2001; 2001US-00984130.
 XX
 XX 28-OCT-1998; 98US-0105971P.
 PR 27-OCT-1999; 99WO-US025031.
 PR 19-APR-2000; 2000US-0198407P.
 PR 30-OCT-2000; 2000US-0243792P.
 PR 18-APR-2001; 2001US-00836353.
 XX
 XX (NIJU/) NI J.
 PA (YOUN/) YOUNG P E.
 PA (KENN/) KENNY J J.
 PA (OLSE/) OLSEN H S.
 PA (MOOR/) MOORE P A.
 PA (WEIY/) WEI Y.
 PA (GREE/) GREENE J M.
 PA (RUBE/) RUBEN S M.
 PA (LIUD/) LIU D.
 PA (CROC/) CROCKER P R.
 XX
 XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
 PI Ruben SM, Liu D, Crocker PR;
 XX WPI; 2003-567103/53.
 XX
 XX New human secreted nucleic acid molecules and polypeptides, useful for
 PT preventing, treating, or ameliorating a medical condition, such as
 PT cancer, inflammation, immune disorders, neurological and blood clotting
 PT disorders.
 XX
 XX Disclosure; Page 343; 454pp; English.
 XX
 XX The invention relates to an isolated nucleic molecule that is at least
 CC 95% identical to 18 human cDNA sequences representing 12 novel genes
 CC encoding secreted proteins or a polynucleotide fragment of the cDNA
 CC sequence contained in American Type Culture Collection (ATCC) deposit No.
 CC defined in the specification, its species homologue, a variant or allelic
 CC variant of the polynucleotide having a polynucleotide capable of
 CC hybridizing under conditions the polynucleotide, where the polynucleotide
 CC does not hybridize under stringent conditions to a nucleic acid molecule
 CC having a nucleotide sequence of only A or T residues. Also included are
 CC recombinant vectors, host cells (for producing the polypeptide), the
 CC secreted polypeptide (comprising a sequence that is at least 95%
 CC identical to a polypeptide fragment, domain, epitope, full-length
 CC protein, variant, allelic variant or species homologue), antibodies that
 CC specifically bind to the polypeptides, diagnosing, treating, preventing
 CC or ameliorating a medical condition by administering the polynucleotide
 CC or the polypeptide, the gene corresponding to the cDNA sequence and
 CC identifying an activity in a biological assay (by expressing the cDNA
 CC sequence in a cell, isolating the supernatant, and detecting an activity
 CC in a biological assay and identifying the protein in the supernatant
 CC having the activity). The polypeptides, nucleic acids and antibodies are
 CC useful for diagnosing a pathological condition or a susceptibility to a
 CC pathological condition, for preventing, treating, or ameliorating a
 CC medical condition, such as cancer, inflammation and other immune
 CC disorders, neurological and blood clotting disorders (many examples are
 CC given in the specification). The nucleic acids are also useful for
 CC chromosome identification, radiation hybrid mapping or long-range
 CC restriction mapping. The polypeptides and antibodies are useful for

CC providing immunological probes for differential identification of the
CC tissues immunohistochemistry assays. The polypeptide, polynucleotide,
CC agonist or antagonist may also be used as a food additive or preservative
CC to increase or decrease storage capabilities, fat content or other
CC nutritional components. The present is a peptide motif from a secreted
CC protein of the invention.
XX
SQ Sequence 33 AA;

Query Match 2.8%; Score 33; DB 6; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.6e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 TNGYQKTGVDYKCPVHGNCTKLNLRVLSNV 96
DB 1 TNGYQKTGVDYKCPVHGNCTKLNLRVLSNV 33

RESULT 38
ADE86651
ID ADE86651 standard; peptide; 33 AA.
AC ADE86651;
XX
XX 29-JAN-2004 (first entry)
DT Transmembrane domain.
DE
XX human; secreted protein; cancer; liver disorder; hepatitis;
KW neural disorder; Alzheimer's disease; transmembrane domain.
XX
XX Homo sapiens.
OS
XX US2003129685-A1.
PN
XX 10-JUL-2003.
XX
XX 18-APR-2001; 2001US-00836353.
PF
XX 28-OCT-1998; 98US-0105971P.
PR
XX 27-OCT-1999; 99WO-US025031.
PR
XX 19-APR-2000; 2000US-0198407P.
XX
XX (NIJ/) NI J.
PA (YOUN/) YOUNG P E.
PA (KERN/) KENNY J J.
PA (OLSE/) OLSEN H S.
PA (MOOR/) MOORE P A.
PA (WEIV/) WEI Y.
PA (GREE/) GREENE J M.
PA (RUBE/) RUBEN S M.

XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
PI Ruben SM;
PI
XX WPI; 2004-020335/02.
XX
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX
XX Disclosure; SEQ ID NO 102; 380pp; English.

XX The invention relates to an isolated nucleic acid sequence, or its
CC allelic variant, a fragment of the cDNA sequence, or its fragment,
CC domain, epitope or species homologue. The nucleic acid is useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition e.g., cancer, liver disorders such as hepatitis or neural
CC disorders such as Alzheimer's disease. The present sequence represents
CC the amino acid sequence of a transmembrane domain.
XX
SQ Sequence 33 AA;

Query Match 2.8%; Score 33; DB 8; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.6e-24;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 64 TNGYQKTGVDYKCPVHGNCTKLNLRVLSNV 96
DB 1 TNGYQKTGVDYKCPVHGNCTKLNLRVLSNV 33

RESULT 39
AAB30927
ID AAB30927 standard; peptide; 24 AA.
XX
XX AAB30927;
AC
XX 02-APR-2001 (first entry)
DT
DE Peptide derived from a human alpha1 integrin chain.
XX
XX Human; integrin; alpha1 subunit; fibroblast; muscle cell; chondrocyte;
KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;
KW osteoporosis; cartilage damage; bone damage; cartilage.
XX
XX Homo sapiens.
OS
XX WO200075187-A1.
FN
XX 14-DEC-2000.
PD
XX 31-MAY-2000; 2000WO-SE001135.
PF
XX 03-JUN-1999; 99SE-00002056.
PR
XX (ACTI-) ACTIVE BIOTECH AB.
PA
XX Gullberg D;
PI
XX WPI; 2001-071061/08.

XX Integrin subunit alpha 11 or integrin heterodimer comprising subunit
PT alpha 11 in association with subunit beta, useful for treating muscle
PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.
XX
XX Claim 23; Page 40; 79pp; English.

XX The present sequence is derived from the cytoplasmic domain of the human
CC integrin subunit, designated alpha11. The alpha11 polynucleotide and
CC polypeptide are useful as markers of cell target molecules, such as
CC fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally
CC derived cells or stem cells. They are also used for determining the
CC differential stage of cells during differentiation, development in
CC pathological conditions, in tissue regeneration, in transplantation or in
CC therapeutic and physiological repair of tissues. The pathological
CC conditions involving subunit alpha11 are selected from damage of cells,
CC muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,
CC osteoarthritis and osteoporosis, damage of cartilage and bone, and
CC cartilage and bone diseases. The polypeptide is useful for detecting the
CC formation of cartilage during embryonic development, for detecting
CC physiological therapeutic repair of cartilage and muscle, for selection
CC and analysis, or for sorting, isolating or purification of chondrocytes
CC and muscle cells, for detecting regeneration of cartilage or chondrocytes
CC during transplantation of cartilage or chondrocytes during
CC transplantation of cartilage or chondrocytes, respectively, or of muscle
CC or muscle cells during transplantation of muscle or muscle cells,
CC respectively, and for studies of differentiation or chondrocytes or
CC muscle cells
XX
SQ Sequence 24 AA;

Query Match 2.0%; Score 24; DB 4; Length 24;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1165 KLGFFRSARRRRPGLDPTPKVLE 1188
 Db 1 KLGFFRSARRRRPGLDPTPKVLE 24

RESULT 40

AAAY76112
 ID AAY76112 standard; protein; 158 AA.
 XX AC AAY76112;
 XX AC AAY76112;

DT 27-MAR-2000 (first entry)

DE Rat integrin homologue, SEQ ID NO:391.

XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KW secreted; transmembrane; inflammation; cancer; neurological disease;
 KW angiogenesis; tumour vascularisation; growth disorder;
 KW developmental disorder; skin wound; hair follicle disorder;
 KW anti-inflammatory; cytostatic; neuroprotective; vulnery.

XX Rattus sp.

XX WO9955865-A1.

XX 04-NOV-1999.

XX 29-APR-1999; 99WO-NZ0000051.

XX 29-APR-1998; 98US-00069726.

XX 09-NOV-1998; 98US-00188930.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Strachan L, Sleeman M, Watson JD, Orrust R, Kumble A, Murison JG;
 PI N-PSDB; AAZ61820.

XX WPI; 2000-072177/06.

XX N-PSDB; AAZ61820.

XX Novel polynucleotides useful for the treatment of various conditions
 PT including wounds and cancer.

XX Claim 4; Page 221; 235pp; English.

XX The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of cancer
 CC cells, to modulate angiogenesis and tumour vascularisation, to modulate
 CC skin inflammation, to modulate epithelial cell growth and to inhibit
 CC binding of HIV-1 to leukocytes. The invention may also be used to treat
 CC growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded by
 CC cDNA sequences derived from several mouse, rat or human skin cell types.
 CC Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and AAY76119
 CC are proteins with an N-terminal signal sequence, indicating that they are
 CC secreted. Sequences AAY75985-Y75989, AAY76061-Y76071, AAY76106-Y76109 and
 CC AAY76121-Y76122 are proteins with one or more putative transmembrane
 CC domains

XX Sequence 158 AA;

Query Match 1.8%; Score 21; DB 3; Length 158;

Best Local Similarity 100.0%; Pred. No. 8.1e-12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 RPVVQINASLHFPSPKINIFH 656

Db 57 RPVVQINASLHFPSPKINIFH 77

RESULT 41

AAAB56051
 ID AAB56051 standard; protein; 158 AA.

XX AC AAB56051;

XX 08-MAR-2001 (first entry)

XX Skin cell protein, SEQ ID NO: 391.

XX Rat; skin cell; cytostatic; antiinflammatory; anti-HIV; nootropic;
 KW neuroprotective; vulnery; immunomodulatory; vaccine;
 KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 KW inflammation; neurological disease.

XX Rattus sp.

XX WO200069884-A2.

XX 23-NOV-2000.

XX 15-MAY-2000; 2000WO-NZ0000075.

XX 14-MAY-1999; 99US-00312283.

XX (GENE-) GENESIS RES & DEV CORP LTD.

XX Watson JD, Strachan L, Orrust R, Sleeman M, Kumble KD;
 PI Murison JG;

XX WPI; 2001-007495/01.

XX N-PSDB; AAC99753.

XX New isolated polynucleotide used in the identification of genetic
 PT disorders and encoding polypeptides used for treating inflammatory
 PT disease, cancer and neurological diseases.

XX Claim 4; Page 296-297; 352pp; English.

XX The present sequence is a polypeptide which is expressed in mammalian
 CC skin cells. The polypeptide is useful for stimulating keratinocyte growth
 CC and motility, inhibiting the growth of cancer cells, modulating
 CC angiogenesis, inhibiting angiogenesis and vascularisation of tumours,
 CC modulating skin inflammation, stimulating the growth of epithelial cells,
 CC inhibiting the binding of human immunodeficiency virus (HIV)-1 to
 CC leukocytes, and treating inflammatory disease, cancer and neurological
 CC diseases. The polynucleotide can be used as a marker, in the
 CC identification of genetic disorders, and for the design of
 CC oligonucleotides for examining expression patterns

XX Sequence 158 AA;

Query Match 1.8%; Score 21; DB 4; Length 158;

Best Local Similarity 100.0%; Pred. No. 8.1e-12;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 636 RPVVQINASLHFPSPKINIFH 656

Db 57 RPVVQINASLHFPSPKINIFH 77

RESULT 42

ABBY72251
 ID ABBY72251 standard; protein; 158 AA.

XX AC ABBY72251;

XX 04-APR-2002 (first entry)

XX Rat protein isolated from skin cells SEQ ID NO: 391.

KW Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
 KW developmental defect; inflammatory disease; dermatological; vulnary;
 KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.
 XX
 OS Rattus sp.
 XX
 FN WO200190357-A1.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-NZ000099.
 XX
 PF 24-MAY-2000; 2000US-0206650P.
 PR
 PR 25-JUL-2000; 2000US-0221232P.
 XX
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA
 XX Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;
 PI Kumble KD;
 PI
 XX WPI; 2002-122020/16.
 DR
 XX New polynucleotides and polypeptides encoded by the polynucleotides
 PT isolated from skin cells, useful for treating skin wounds, cancers,
 PT growth and developmental defects, inflammatory diseases, or for
 PT modulating immune responses.
 XX
 XX Example 2; Page 246; 466pp; English.
 PS
 XX The present invention provides the protein and coding sequences of cDNAs
 CC isolated from human, murine and rat skin cell libraries. The sequences
 CC can be used in the development of therapeutic agents useful in the
 CC treatment of skin diseases, including skin wounds, cancer, growth
 CC defects, developmental defects and inflammatory diseases. The proteins
 CC have important roles in the induction of hair growth, cell proliferation
 CC and cell-cell interaction, in maintaining tissue integrity, in wound
 CC healing and in modulating immune responses. The present sequence is a
 CC polypeptide of the invention
 XX
 XX Sequence 158 AA;
 SQ
 Query Match 1.8%; Score 21; DB 5; Length 158;
 Best Local Similarity 100.0%; Pred. No. 8.1e-12;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 636 RPVVQINASLHFPSPKINIFH 656
 DB 57 RPVVQINASLHFPSPKINIFH 77
 RESULT 43
 AAY76111
 ID AAY76111 standard; protein; 85 AA.
 XX
 AC AAY76111;
 XX
 XX 27-MAR-2000 (first entry)
 DT
 XX Rat integrin homologue, SEQ ID NO:390.
 DE
 XX Skin; dermal papilla; keratinocyte; neonatal foreskin fibroblast;
 KW embryonic skin cell; keratinocyte stem cell; transit amplifying cell;
 KW secreted; transmembrane; inflammation; cancer; neurological disease;
 KW angiogenesis; tumour vascularisation; growth disorder;
 KW developmental disorder; skin wound; hair follicle disorder;
 KW anti-inflammatory; cytostatic; neuroprotective; vulnary.
 XX
 OS Rattus sp.
 XX
 XX WO9955865-A1.
 FN
 XX 04-NOV-1999.
 PD
 XX (GENE-) GENESIS RES & DEV CORP LTD.

PF 29-APR-1999; 99WO-NZ000051.
 XX
 PR 29-APR-1998; 98US-00069726.
 PR 09-NOV-1998; 98US-00189930.
 XX
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA
 XX Strachan L, Sleeman M, Watson JD, Onrust R, Kumble A, Murison JG;
 PI WPI; 2000-072177/06.
 DR N-PSDB; AAZ61819.
 DR
 XX Novel polynucleotides useful for the treatment of various conditions
 PT including wounds and cancer.
 PT
 XX Claim 4; Page 221; 235pp; English.
 PS
 XX The invention relates to novel nucleic acid sequences derived from rat
 CC dermal papilla, human keratinocytes and neonatal foreskin fibroblasts,
 CC and mouse embryonic skin, keratinocyte stem cells and transit amplifying
 CC cells. Polypeptides of the invention may be used to treat inflammation,
 CC cancer and neurological diseases. The proteins may be used to stimulate
 CC the growth and motility of keratinocytes, to inhibit the growth of cancer
 CC cells, to modulate angiogenesis and tumour vascularisation, to modulate
 CC skin inflammation, to modulate epithelial cell growth and to inhibit
 CC binding of HIV-1 to leukocytes. The invention may also be used to treat
 CC growth and developmental defects, skin wounds and hair follicle
 CC disorders. Sequences AAY75942-Y76123 represent polypeptides encoded by
 CC cDNA sequences derived from several mouse, rat or human skin cell types.
 CC Sequences AAY75942-Y75947, AAY76020-Y76021, AAY76094-Y76104 and AAY76119
 CC are proteins with an N-terminal signal sequence, indicating that they are
 CC secreted. Sequences AAY75986-Y75989, AAY76061-Y76071, AAY76106-Y76109 and
 CC AAY76121-Y76122 are proteins with one or more putative transmembrane
 CC domains
 XX
 XX Sequence 85 AA;
 SQ
 Query Match 1.4%; Score 17; DB 3; Length 85;
 Best Local Similarity 100.0%; Pred. No. 4.4e-08;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 528 NGTLKDSHSYONARFGS 544
 DB 28 NGTLKDSHSYONARFGS 44
 RESULT 44
 AAB56050
 ID AAB56050 standard; protein; 85 AA.
 XX
 AC AAB56050;
 XX
 XX 08-MAR-2001 (first entry)
 DT
 XX Skin cell protein, SEQ ID NO: 390.
 DE
 XX Rat; skin cell; cytostatic; antiinflammatory; anti-HIV; nootropic;
 KW neuroprotective; vulnary; immunomodulatory; vaccine; inhibition;
 KW keratinocyte growth stimulation; cancer; angiogenesis inhibition;
 KW inflammation; neurological disease.
 XX
 OS Rattus sp.
 XX
 XX WO200069984-A2.
 FN
 XX 23-NOV-2000.
 PD
 XX 15-MAY-2000; 2000WO-NZ000075.
 PF
 XX 14-MAY-1999; 99US-00312283.
 PR
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA
 XX

PI Watson JD, Strachan L, Onrust R, Sleeman M, Kumble KD;
PI Murison JG;
XX
XX
XX WPI: 2001-007495/01.
XX N-PSDB; AAC99752.
XX
XX New isolated polynucleotide used in the identification of genetic
PT disorders and encoding polypeptides used for treating inflammatory
PT disease, cancer and neurological diseases.
XX
XX Claim 4; Page 296; 352pp; English.
XX
XX The present sequence is a polypeptide which is expressed in mammalian
CC skin cells. The polypeptide is useful for stimulating keratinocyte growth
CC and motility, inhibiting the growth of cancer cells, modulating
CC angiogenesis, inhibiting angiogenesis and vascularisation of tumours,
CC modulating skin inflammation, stimulating the growth of epithelial cells,
CC inhibiting the binding of human immunodeficiency virus (HIV)-1 to
CC leukocytes, and treating inflammatory disease, cancer and neurological
CC diseases. The polynucleotide can be used as a marker, in the
CC identification of genetic disorders, and for the design of
CC oligonucleotides for examining expression patterns
XX
XX Sequence 85 AA;
SQ

Query Match 1.4%; Score 17; DB 4; Length 85;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 528 NGTLKDSHSYQNARFGS 544
Db 28 NGTLKDSHSYQNARFGS 44

RESULT 45
ABB72250
ID ABB72250 standard; protein; 85 AA.
XX
XX ABB72250;
XX
XX 04-APR-2002 (first entry)
XX
XX Rat protein isolated from skin cells SEQ ID NO: 390.
DE
DE Human; rat; mouse; skin cell; skin wound; cancer; growth defect;
KW developmental defect; inflammatory disease; dermatological; vulnary;
KW immunomodulator; anti-inflammatory; cytostatic; neuroprotective.
XX
XX Rattus sp.
XX
XX WO200190357-A1.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001WO-NZ000099.
XX
XX 24-MAY-2000; 2000US-0206650P.
XX
XX 25-JUL-2000; 2000US-0221232P.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX
XX Watson JD, Strachan L, Sleeman M, Onrust R, Murison JG;
PI Kumble KD;
PI
XX
XX WPI: 2002-122020/16.
XX
XX New polynucleotides and polypeptides encoded by the polynucleotides
PT isolated from skin cells, useful for treating skin wounds, cancers,
PT growth and developmental defects, inflammatory diseases, or for
PT modulating immune responses.
XX
XX Example 2; Page 245; 466pp; English.
PS
XX

CC The present invention provides the protein and coding sequences of cDNAs
CC isolated from human, murine and rat skin cell libraries. The sequences
CC can be used in the development of therapeutic agents useful in the
CC treatment of skin diseases, including skin wounds, cancer, growth
CC defects, developmental defects and inflammatory diseases. The proteins
CC have important roles in the induction of hair growth, cell proliferation
CC and cell-cell interaction, in maintaining tissue integrity, in wound
CC healing and in modulating immune responses. The present sequence is a
XX polypeptide of the invention
XX Sequence 85 AA;
SQ

Query Match 1.4%; Score 17; DB 5; Length 85;
Best Local Similarity 100.0%; Pred. No. 4.4e-08;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 528 NGTLKDSHSYQNARFGS 544
Db 28 NGTLKDSHSYQNARFGS 44

RESULT 46
AAU19634
ID AAU19634 standard; protein; 148 AA.
XX
XX AAU19634;
XX
XX 04-DEC-2001 (first entry)
XX
XX Human novel extracellular matrix protein, Seq ID No 284.
XX
XX Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;
KW antianemic; antirheumatic; antisclerotic; cardiac; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmic; cytostatic;
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disorder; breast neoplasia; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW cardiac arrest; tachycardia; angina; infection; corneal infections;
KW wound healing; immunogen; gene therapy; antisense; food additive.
XX
XX Homo sapiens.
XX
XX WO200153368-A1.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001348.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX
XX 04-FEB-2000; 2000US-0180628P.
XX
XX 24-FEB-2000; 2000US-0184664P.
XX
XX 02-MAR-2000; 2000US-0186350P.
XX
XX 16-MAR-2000; 2000US-0189874P.
XX
XX 17-MAR-2000; 2000US-0190076P.
XX
XX 18-APR-2000; 2000US-0198123P.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX 07-JUN-2000; 2000US-0209467P.
XX
XX 28-JUN-2000; 2000US-0214886P.
XX
XX 30-JUN-2000; 2000US-0215135P.
XX
XX 07-JUL-2000; 2000US-0216647P.
XX
XX 07-JUL-2000; 2000US-0216680P.
XX
XX 11-JUL-2000; 2000US-0217487P.
XX
XX 14-JUL-2000; 2000US-0217496P.
XX
XX 26-JUL-2000; 2000US-0220963P.
XX
XX 26-JUL-2000; 2000US-0220964P.
XX
XX 14-AUG-2000; 2000US-0224518P.
XX
XX 14-AUG-2000; 2000US-0224519P.
XX
XX 14-AUG-2000; 2000US-0225213P.
XX
XX 14-AUG-2000; 2000US-0225214P.
XX
XX 14-AUG-2000; 2000US-0225266P.

[illegible]

CC /cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),
CC infections caused by bacteria, viruses and fungi and ocular disorders
CC (e.g. corneal infections). Other uses include wound healing, maintenance
CC of organs before transplantation, support of cell culture of primary

Query Match 1.38; Score 15; DB 4; Length 148;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGSGNSIYFW 178
Db 47 DIVVLGSGNSIYFW 61

RESULT 47
AAU19794
ID AAU19794 standard; protein; 148 AA.

XX AC AAU19794;

XX DT 04-DEC-2001 (first entry)

XX DE Human novel extracellular matrix protein, Seq ID No 444.

XX KW Human; secreted extracellular matrix protein; immunomodulatory; Anti-HIV;
KW antianemic; antirheumatic; antisclerotic; cardiant; vascular;
KW cerebroprotective; thrombolytic; antimicrobial; opthalmic; cytostatic;
KW antialzheimers; immune/autoimmune disease; HIV infection; anaemia;
KW human immunodeficiency virus; rheumatoid arthritis; multiple sclerosis;
KW cancers; hyperproliferative disorder; breast neoplasm; melanoma;
KW Sezary syndrome; Gaucher's disease; neurological diseases;
KW Alzheimer's disease; Parkinson's disease; cardiovascular disorder;
KW Cardiac arrest; tachycardia; angina; infection; corneal infections;
KW wound healing; immunogen; Gene therapy; antisense; food additive.

XX OS Homo sapiens.

XX PN WC200155368-A1.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001348.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 30-JUN-2000; 2000US-0214886P.

XX PR 07-JUL-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0215647P.

XX PR 11-JUL-2000; 2000US-0216880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 14-JUL-2000; 2000US-0217496P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 14-AUG-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 12-SEP-2000; 2000US-0232081P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 21-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 26-SEP-2000; 2000US-0234998P.
PR 27-SEP-2000; 2000US-0235844P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0241826P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0248526P.
PR 08-NOV-2000; 2000US-0248527P.
PR 08-NOV-2000; 2000US-0248528P.
PR 08-NOV-2000; 2000US-0248532P.
PR 08-NOV-2000; 2000US-0248609P.
PR 08-NOV-2000; 2000US-0248610P.
PR 08-NOV-2000; 2000US-0248611P.

PR	08-NOV-2000;	2000US-0246613P.	Matches	15;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
PR	17-NOV-2000;	2000US-0249207P.	Qy	164	DIVILDGNSIYPW	178						
PR	17-NOV-2000;	2000US-0249208P.										
PR	17-NOV-2000;	2000US-0249209P.										
PR	17-NOV-2000;	2000US-0249210P.	Db	47	DIVILDGNSIYPW	61						
PR	17-NOV-2000;	2000US-0249211P.										
PR	17-NOV-2000;	2000US-0249212P.										
PR	17-NOV-2000;	2000US-0249213P.										
PR	17-NOV-2000;	2000US-0249214P.										
PR	17-NOV-2000;	2000US-0249215P.										
PR	17-NOV-2000;	2000US-0249216P.										
PR	17-NOV-2000;	2000US-0249217P.										
PR	17-NOV-2000;	2000US-0249218P.										
PR	17-NOV-2000;	2000US-0249244P.										
PR	17-NOV-2000;	2000US-0249245P.										
PR	17-NOV-2000;	2000US-0249264P.										
PR	17-NOV-2000;	2000US-0249265P.										
PR	17-NOV-2000;	2000US-0249297P.										
PR	17-NOV-2000;	2000US-0249299P.										
PR	17-NOV-2000;	2000US-0249300P.										
PR	01-DEC-2000;	2000US-0250160P.										
PR	01-DEC-2000;	2000US-0250391P.										
PR	05-DEC-2000;	2000US-0251030P.										
PR	05-DEC-2000;	2000US-0251988P.										
PR	05-DEC-2000;	2000US-0256719P.										
PR	06-DEC-2000;	2000US-0251479P.										
PR	08-DEC-2000;	2000US-0251856P.										
PR	08-DEC-2000;	2000US-0251868P.										
PR	08-DEC-2000;	2000US-0251869P.										
PR	08-DEC-2000;	2000US-0251989P.										
PR	08-DEC-2000;	2000US-0251990P.										
PR	11-DEC-2000;	2000US-0254097P.										
PR	05-JAN-2001;	2001US-0259678P.										
XX												
PA	(HUMA-) HUMAN GENOME SCI INC.											
XX												
PI	Rosen CA, Barash SC, Ruben SM;											
XX												
XX												
DR	WT1; 2001-455572/50.											
DR	N-PSDB; AAS31365.											
XX												
PT	Nucleic acid molecules encoding human secreted extracellular matrix											
PT	proteins, used in preventing, treating or ameliorating a disorder, e.g.											
PT	Alzheimer's and Parkinson's diseases and cancers.											
XX												
PS	Claim 11; SEQ ID NO 444; 577pp; English.											
XX												
CC	The invention relates to isolated nucleic acid molecules encoding novel											
CC	human secreted extracellular matrix proteins (SPs). The polynucleotides											
CC	and proteins are used to prevent, treat a medical condition in e.g.											
CC	humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep. For											
CC	example, disorders associated with decreased expression of SPs. The SP											
CC	polynucleotide or a vector expressing them may be administered to treat											
CC	diseases by gene therapy. Antisense molecules may be administered to down											
CC	regulate expression of SPs by binding with the cells own genes and											
CC	preventing their expression. The polynucleotides may also be used as DNA											
CC	probes in diagnostic assays. The SPs may also be used as antigens to											
CC	produce antibodies and to identify modulators (agonists and antagonists)											
CC	of the SPs. The anti-(SP) antibodies and antagonists may also be used to											
CC	down regulate expression and activity of SP and as diagnostic agents for											
CC	detecting the presence of SPs in samples. The disorders include for											
CC	example: immune/autoimmune diseases (e.g. HIV (human immunodeficiency											
CC	virus) infections, anaemia, rheumatoid arthritis and multiple sclerosis),											
CC	cancers and hyperproliferative disorders (e.g. melanomas, neoplasms of											
CC	the breast or liver, Sezary syndrome and Gaucher's disease), neurological											
CC	diseases (e.g. Alzheimer's disease, Parkinson's disease) cardio-											
CC	/cerebrovascular disorders (e.g. cardiac arrest, tachycardia and angina),											
CC	infections caused by bacteria, viruses and fungi and ocular disorders											
CC	(e.g. corneal infections). Other uses include wound healing, maintenance											
CC	of organs before transplantation, support of cell culture of primary											

Query Match 1.3%; Score 15; DB 4; Length 148;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;


```

XX CC The invention relates to novel genes (ABQ66521-ABQ66785) and proteins
XX CC (ABP47846-ABP48110) useful for preventing, treating or ameliorating
XX CC medical conditions e.g. by protein or gene therapy. The genes are
XX CC isolated from a range of human tissues disclosed in the specification.
XX CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX CC anaemia, autoimmune rheumatoid arthritis, diabetes mellitus, Crohn's disease,
XX CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX CC infectious diseases such as viral, bacterial, fungal and parasitic
XX CC infections. Note: The sequence data for this patent did not form part of
XX CC the printed specification, but was obtained in electronic format directly
XX CC from USPTO at seqdata.uspto.gov/sequence.html?DocId=99909764870
XX SQ
SQ Sequence 148 AA;
Query Match 1.3%; Score 15; DB 5; Length 148;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 DIVIVLDGNSIYPW 178
Db 47 DIVIVLDGNSIYPW 61
|||||
RESULT 50
ADCL0816
ID ADC10816 standard; protein; 148 AA.
XX
AC ADC10816;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human extracellular matrix protein from gene 9.
XX
KW Extracellular matrix protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal; cardiac;
KW cardiovascular; nephrotropic; antiinflammatory; muscular;
KW respiratory; immunosuppressive; cerebroprotective; vasotropic;
KW neotrophic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; Human.
XX
OS Homo sapiens.
XX
PN US2003059875-A1.
XX
PD 27-MAR-2003.
XX
PF 19-APR-2002; 2002US-00125540.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214896P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 18-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226888P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232357P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244828P.
PR 01-NOV-2000; 2000US-0244829P.

```

PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251031P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256179P.
PR 08-DEC-2000; 2000US-0251799P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2000US-0254037P.
PR 17-JAN-2001; 2001US-00764870.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-743765/70.
XX N-PSDB; ADCI0551.
XX
XX New isolated nucleic acids and polypeptides, useful for diagnosing,
XX treating, and/or preventing disorders, such as cancer, infections,
XX cardiovascular and inflammatory diseases.
XX
XX Claim 11; SEQ ID NO 284; 235pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (cDNA)
XX encoding a human extracellular matrix protein, representing one of 161
XX novel genes. Also included are recombinant vectors, host cells
XX (expressing the protein), the extracellular matrix proteins (including
XX their fragments, epitopes and homologues), an isolated antibody that
XX binds specifically to the protein, diagnosing a pathological condition or
XX susceptibility to a pathological condition (comprising determining the
XX presence or absence of a mutation in the nucleic acid and diagnosing a
XX condition based on the presence or absence of the mutation), diagnosing a
XX pathological condition or susceptibility to a pathological condition

CC (comprising determining the presence or amount of expression of the
CC protein in a biological sample and diagnosing a condition based on the
CC presence or amount of expression of the protein), preventing, treating or
CC ameliorating a medical condition by administering the nucleic acid or
CC protein to a mammalian subject, identifying a binding partner to the
CC protein, the gene corresponding to the cDNA sequence, and identifying an
CC activity in a biological assay (comprising expressing the nucleic acid in
CC a cell, isolating the supernatant, detecting an activity in a biological
CC assay and identifying the protein in the supernatant having the
CC activity). The nucleic acids and proteins display the following
CC activities: Cytostatic, antibacterial, Virucide, Neuroprotective,
CC Gynaecological, Gastrointestinal-Gen, Cardiant, Cardiovascular-Gen,
CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,
Query Match 1.3%; Score 15; DB 7; Length 148;
Best Local Similarity 100.0%; Pred. No. 6.9e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 DIVIVDGSNSIYPW 178
Db 47 DIVIVDGSNSIYPW 61
|||||
|||||
RESULT 51
ADCI0976
ID ADCI0976 standard; protein; 148 AA.
XX
AC ADCI0976;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human protein from extracellular matrix gene 9 #2.
XX
KW Extracellular matrix protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW notropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; Human.
XX
OS Homo sapiens.
XX
XX US2003059875-A1.
XX
XX 27-MAR-2003.
XX
XX 19-APR-2002; 2002US-00125540.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216800P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0228868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249298P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764870.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-743765/70.
DR N-PSDB; ADC10711.
XX
XX New isolated nucleic acids and polypeptides, useful for diagnosing,
PT treating, and/or preventing disorders, such as cancer, infections,
PT cardiovascular and inflammatory diseases.
XX
XX Claim 11; SEQ ID NO 444; 235pp; English.
PS
XX
XX The invention relates to an isolated nucleic acid molecule (cDNA)
CC encoding a human extracellular matrix protein, representing one of 161
CC novel genes. Also included are recombinant vectors, host cells
CC (expressing the protein), the extracellular matrix proteins (including
CC their fragments, epitopes and homologues), an isolated antibody that
CC binds specifically to the protein, diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or absence of a mutation in the nucleic acid and diagnosing a
CC condition based on the presence or absence of the mutation), diagnosing a
CC pathological condition or susceptibility to a pathological condition
CC (comprising determining the presence or amount of expression of the
CC protein in a biological sample and diagnosing a condition based on the
CC presence or amount of expression of the protein), preventing, treating or
CC ameliorating a medical condition by administering the nucleic acid or
CC protein to a mammalian subject, identifying a binding partner to the
CC protein, the gene corresponding to the cDNA sequence, and identifying an
CC activity in a biological assay (comprising expressing the nucleic acid in

CC a cell, isolating the supernatant, detecting an activity in a biological
 CC assay and identifying the protein in the supernatant having the
 CC activity). The nucleic acids and proteins display the following
 CC activities Cytostatic, antibacterial, Virucide, Neuroprotective,
 CC Gynaecological, Gastrointestinal-Gen, Cardiac, Cardiovascular-Gen,
 CC Nephrotropic, Antiinflammatory, Muscular-Gen, Respiratory-Gen,

Query Match 1.3%; Score 15; DB 7; Length 148;
 Best Local Similarity 100.0%; Pred. No. 6.9e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLQGSNSIYPW 178
 Db 47 DIVIVLQGSNSIYPW 51

RESULT 52
 AAE33535
 ID AAE33535 standard; protein; 192 AA.

XX AC AAE33535;

XX DT 23-OCT-2003 (revised)

XX DT 02-APR-2003 (first entry)

XX DT Rat-human alpha 1-I domain fusion protein, RdeltaH.

XX KW Human; very late activation antigen; VLA-1; betal containing integrin;
 KW immunological disorder; inflammatory disorder; skin related condition;
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarteritis nodosa;
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;
 KW graft versus host disease; myocardial ischaemia; fusion protein; rat.

XX OS Homo sapiens.

XX OS Rattus sp.

XX OS Chimeric.

XX PN WO200283854-A2.

XX PD 24-OCT-2002.

XX PF 12-APR-2002; 2002WO-US011521.

XX PR 13-APR-2001; 2001US-0283794P.

XX PR 06-JUL-2001; 2001US-0303689P.

XX PA (BIOJ) BIOGEN INC.

XX PI Lyne PD, Garber EA, Saldanha JW, Karpusas M;

XX DR WPI; 2003-093009/08.

XX DR New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-

PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or
 PT fibrosis.

XX PS Example 24; Page 93; 248pp; English.

XX CC The present invention relates to novel antibodies that specifically bind
 CC to very late activation (VLA-1; betal containing integrins) antigens and
 CC methods of using these antibodies to treat immunological disorders. The
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 CC mediated immunological or inflammatory disorders such as skin related
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal

CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),
 CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,
 CC periarteritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or
 CC immediate hypersensitivity), graft and transplant rejections, graft
 CC versus host disease, conjunctivitis, swelling occurring after injury,
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is
 CC rat-human alpha 1-I domain fusion protein, RdeltaH. This sequence is used
 CC in the exemplification of the invention. (Updated on 23-OCT-2003 to
 CC standardise OS field)

XX SQ Sequence 192 AA;

Query Match 1.3%; Score 15; DB 6; Length 192;

Best Local Similarity 100.0%; Pred. No. 8.7e-06;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVIVLQGSNSIYPW 178

Db 4 DIVIVLQGSNSIYPW 18

RESULT 53

AAE33536

ID AAE33536 standard; protein; 192 AA.

XX AC AAE33536;

XX DT 02-APR-2003 (first entry)

XX DT Rat alpha 1-I domain protein #1.

XX KW Rat; very late activation antigen; VLA-1; betal containing integrin;
 KW immunological disorder; inflammatory disorder; skin related condition;
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarteritis nodosa;
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;
 KW graft versus host disease; myocardial ischaemia, alpha 1-I protein.

XX OS Rattus sp.

XX PN WO200283854-A2.

XX PD 24-OCT-2002.

XX PF 12-APR-2002; 2002WO-US011521.

XX PR 13-APR-2001; 2001US-0283794P.

XX PR 06-JUL-2001; 2001US-0303689P.

XX PA (BIOJ) BIOGEN INC.

XX PI Lyne PD, Garber EA, Saldanha JW, Karpusas M;

XX DR WPI; 2003-093009/08.

XX DR New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-

PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or

PT fibrosis.

PS Example 24; Page 93-94; 248pp; English.

XX

CC The present invention relates to novel antibodies that specifically bind to very late activation (VLA-1; beta1 containing integrins) antigens and methods of using these antibodies to treat immunological disorders. The anti-VLA-1 antibodies are useful for preventing or treating VLA-1-mediated immunological or inflammatory disorders such as skin related conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal proliferation of hair follicle cells), fibrosis (e.g. kidney or lung fibrosis), allergic rhinitis, respiratory distress syndrome, asthma, bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease, gastritis, irritable bowel syndrome, colitis and colorectal cancer), vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia, periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis, autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid arthritis, systemic lupus erythematosus and multiple sclerosis), renal failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis, polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or immediate hypersensitivity), graft and transplant rejections, graft versus host disease, conjunctivitis, swelling occurring after injury, myocardial ischaemia or endotoxin shock syndrome. The present sequence is rat 1-I domain protein. This sequence is used in the exemplification of the invention

XX

SQ Sequence 192 AA;

Query Match 1.3%; Score 15; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
|||||
Db 4 DIVIVLDGNSIYPW 18

RESULT 54
AAE33537

ID AAE33537 standard; protein; 192 AA.

AC AAE33537;

XX

XX 02-APR-2003 (first entry)

DT

DE Human alpha 1-I domain protein #1.

XX

KW Human; very late activation antigen; VLA-1; beta1 containing integrin; immunological disorder; inflammatory disorder; skin related condition; psoriasis; eczema; burn; dermatitis; respiratory distress syndrome; fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis; fever; migraine headache; inflammatory bowel disease; Crohn's disease; irritable bowel syndrome; colitis; colorectal cancer; vascular disease; atherosclerosis; thyroiditis; aplastic anaemia; periarthritis nodosa; gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease; osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis; systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome; renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis; KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis; KW graft versus host disease; myocardial ischaemia; alpha 1-I protein.

XX

OS Homo sapiens.

XX

PN WO200283854-A2.

XX

PD 24-OCT-2002.

XX

PF 12-APR-2002; 2002WO-US011521.

XX

PR 13-APR-2001; 2001US-0283794P.

PR 06-JUL-2001; 2001US-0303689P.

XX

PA (BIOJ) BIOGEN INC.

XX

PI Lyne PD, Garber EA, Saldanha JW, Karpusas M;

FI

DR WPI; 2003-093009/08.

XX

PT New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-mediated immunological or inflammatory disorders, e.g. psoriasis, eczema, burns, dermatitis, and abnormal proliferation of hair follicle cells or fibrosis.

PT

XX

XX Example 24; Page 94; 248pp; English.

XX

CC The present invention relates to novel antibodies that specifically bind to very late activation (VLA-1; beta1 containing integrins) antigens and methods of using these antibodies to treat immunological disorders. The anti-VLA-1 antibodies are useful for preventing or treating VLA-1-mediated immunological or inflammatory disorders such as skin related conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal proliferation of hair follicle cells), fibrosis (e.g. kidney or lung fibrosis), allergic rhinitis, respiratory distress syndrome, asthma, bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease, gastritis, irritable bowel syndrome, colitis and colorectal cancer), vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia, periarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis, autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid arthritis, systemic lupus erythematosus and multiple sclerosis), renal failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis, polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or immediate hypersensitivity), graft and transplant rejections, graft versus host disease, conjunctivitis, swelling occurring after injury, myocardial ischaemia or endotoxin shock syndrome. The present sequence is human 1-I domain protein. This sequence is used in the exemplification of the invention

XX

SQ Sequence 192 AA;

Query Match 1.3%; Score 15; DB 6; Length 192;
Best Local Similarity 100.0%; Pred. No. 8.7e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
|||||
Db 4 DIVIVLDGNSIYPW 18

RESULT 55
AAU76851

ID AAU76851 standard; protein; 195 AA.

XX

AC AAU76851;

XX

DT 21-MAY-2002 (first entry)

XX

DE Human integrin alpha subunit Alpha 1 (CD49a) A domain.

XX

KW Human; integrin alpha subunit; A domain; CD49a; integrin beta subunit; A-like domain; inflammatory disorder; skeletal muscle injury; restenosis; KW ischaemia-reperfusion injury; immune complex; parasitic disease; Alpha 1; KW antiinflammatory; vasotropic; antiparasitic; vulnery; gene therapy.

XX

OS Homo sapiens.

XX

PN WO200209737-A1.

XX

PD 07-FEB-2002.

XX

PF 31-JUL-2001; 2001WO-US023957.

XX

PR 31-JUL-2000; 2000US-0221950P.

PR 11-JAN-2001; 2001US-00758493.

PR 13-MAR-2001; 2001US-00805354.

```

XX PA (GEHO ) GEN HOSPITAL CORP.
XX PI Arnaut AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS Example 2; Fig 5; 55pp; English.
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents the human
XX CC integrin alpha subunit Alpha 1 (CD49a) A domain
XX SQ Sequence 195 AA;

Query Match 1.3%; Score 15; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
Db 6 DIVIVLDGNSIYPW 20
|||||
|

RESULT 56
AAU76860
ID AAU76860 standard; protein; 195 AA.
AC AAU76860;
DT 21-MAY-2002 (first entry)
DE Human integrin alpha subunit Alpha 1 (CD49a) variant A domain.
KW Human; integrin alpha subunit; A domain; CD49a; integrin beta subunit;
KW A-like domain; inflammatory disorder; skeletal muscle injury; restenosis;
KW ischaemia-reperfusion injury; immune complex; parasitic disease; mutant;
KW anti-inflammatory; vasotropic; antiparasitic; vulnary; gene therapy;
KW mutin; Alpha 1.
XX OS Homo sapiens.
XX OS Synthetic.
XX FH Key
XX FT Location/Qualifiers
XX FT Misc-difference 193
XX FT /note= "Wild-type Ile substituted by any other amino
XX FT acid"
XX PN WC200209737-A1.
XX PD 07-FEB-2002.
XX PF 31-JUL-2001; 2001WO-US023957.
XX PR 31-JUL-2000; 2000US-0221950P.
XX PR 11-JAN-2001; 2001US-00758493.
XX PR 13-MAR-2001; 2001US-00805354.

```

```

XX PA (GEHO ) GEN HOSPITAL CORP.
XX PI Arnaut AM, Li R, Xiong J;
XX DR WPI; 2002-188687/24.
XX PT Novel high affinity integrin polypeptide useful for treating restenosis
XX PT and parasitic diseases, comprises all or part of variant integrin alpha
XX PT subunit A domain or variant integrin beta subunit A-like domain.
XX PS Claim 45; Page; 55pp; English.
XX CC The invention relates to a high affinity integrin polypeptide comprising
XX CC all or part of a variant integrin alpha subunit A domain or a variant
XX CC integrin beta subunit A-like domain. The polypeptide, preferably the
XX CC CD11b alpha subunit A domain, where I at residue 332 has been replaced by
XX CC G or A, F at residue 313 and A at residue 320 have been replaced by C, or
XX CC V at residue 315 and A at residue 320 have been replaced by C, is useful
XX CC for determining if a test compound is a candidate compound for binding to
XX CC CD11b or for treating an inflammatory disorder, by contacting a test
XX CC compound with the polypeptide and determining if the test compound binds
XX CC to the polypeptide. The integrin subunits are useful for reducing
XX CC skeletal muscle injury, for treating disorders caused by ischaemia-
XX CC reperfusion injury, immune complexes, restenosis and parasitic diseases,
XX CC to purify variant integrin polypeptide ligands and as bait proteins in
XX CC two-hybrid or three-hybrid assays. This sequence represents a human
XX CC integrin alpha subunit Alpha 1 (CD49a) variant A domain. Note: This
XX CC variant sequence is not featured in the specification but has been
XX CC derived from the wild-type protein shown in AAU76851
XX SQ Sequence 195 AA;

Query Match 1.3%; Score 15; DB 5; Length 195;
Best Local Similarity 100.0%; Pred. No. 8.8e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
Db 6 DIVIVLDGNSIYPW 20
|||||
|

RESULT 57
AAB50041
ID AAB50041 standard; protein; 214 AA.
AC AAB50041;
DT 14-MAR-2001 (first entry)
DE Rat alaph integrin domain.
KW Rat; alaph integrin domain; alaphabeta1 function blocking antibody;
KW inflammatory disorder; rheumatoid arthritis; skin condition; asthma;
KW bronchitis; headache; antipyretic; fever; gastrointestinal;
KW vascular disease; autoimmune; respiratory distress syndrome;
KW endotoxin shock syndrome; atherosclerosis.
XX OS Rattus sp.
XX FN WC2000072881-A1.
XX PD 07-DEC-2000.
XX PF 01-JUN-2000; 2000WO-US015004.
XX PR 01-JUN-1999; 99US-0137038P.
XX PR 29-FEB-2000; 2000US-0185336P.
XX PA (BIOJ ) BIOGEN INC.
XX PI De Fougereolles A, Gotwals P, Lobb R, Koteliensky V;
XX

```

DR WPI; 2001-061448/07.
 XX Use of blocking monoclonal antibody capable of binding to an epitope of
 PT VLA-1 for treating inflammatory disorders, in particular arthritis.
 PT
 PS Example 14; Fig 14; 60pp; English.
 XX

CC The present invention relates to the use of an alphanbetal function
 CC blocking antibody capable of binding an epitope of human alpha1 integrin
 CC domain (VLA-1; see AAB50043). The antibody of the present invention is
 CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,
 CC skin related conditions such as psoriasis, eczema, burns and dermatitis,
 CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the
 CC treatment of pain and headaches, or as an antipyretic, for the treatment
 CC of fever, gastrointestinal conditions such as inflammatory bowel
 CC diseases, Crohn's disease, gastritis and vascular diseases, migraine
 CC headaches, periarthritis nodosa, thyroiditis, aplastic anaemia, Hodgkin's
 CC disease, rheumatic fever, type I diabetes, myasthenia gravis, multiple
 CC sclerosis, sarcoidosis, nephrotic syndrome, myocardial ischaemia,
 CC allergic rhinitis, respiratory distress syndrome, endotoxin shock
 CC syndrome and atherosclerosis. The present sequence is the rat alpha1
 CC integrin domain
 XX
 SQ Sequence 214 AA;

Query Match 1.3%; Score 15; DB 4; Length 214;
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
 |||||
 DB 22 DIVIVLDGNSIYPW 36

RESULT 58
 AAB50042
 ID AAB50042 standard; protein; 214 AA.
 AC AAB50042;
 DT 14-MAR-2001 (first entry)
 XX
 DE Human alpha1 integrin domain.
 XX

Human; alpha1 integrin domain; alphanbetal function blocking antibody;
 KW inflammatory disorder; rheumatoid arthritis; skin condition; asthma;
 KW bronchitis; headache; antipyretic; fever; gastrointestinal;
 KW vascular disease; autoimmune; respiratory distress syndrome;
 KW endotoxin shock syndrome; atherosclerosis.
 XX
 OS Homo sapiens.
 XX
 PN WO200072881-A1.
 XX
 PD 07-DEC-2000.
 XX

PF 01-JUN-2000; 2000WO-US015004.
 XX
 PR 01-JUN-1999; 99US-0137038P.
 PR 29-FEB-2000; 2000US-0185336P.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI De Fougereolles A, Gotwals P, Lobb R, Koteliensky V;
 XX
 XX WPI; 2001-061448/07.
 DR
 XX

PT Use of blocking monoclonal antibody capable of binding to an epitope of
 PT VLA-1 for treating inflammatory disorders, in particular arthritis.
 XX
 PS Example 14; Fig 15; 60pp; English.
 XX
 CC The present invention relates to the use of an alphanbetal function

CC blocking antibody capable of binding an epitope of human alpha1 integrin
 CC domain (VLA-1; see AAB50043). The antibody of the present invention is
 CC useful for treating an inflammatory disorder e.g. rheumatoid arthritis,
 CC skin related conditions such as psoriasis, eczema, burns and dermatitis,
 CC asthma, bronchitis, menstrual cramps, tendinitis, bursitis, and the
 CC treatment of pain and headaches, or as an antipyretic, for the treatment
 CC of fever, gastrointestinal conditions such as inflammatory bowel
 CC diseases, Crohn's disease, gastritis and vascular diseases, migraine
 CC headaches, periarthritis nodosa, thyroiditis, aplastic anaemia, Hodgkin's
 CC disease, rheumatic fever, type I diabetes, myasthenia gravis, multiple
 CC sclerosis, sarcoidosis, nephrotic syndrome, myocardial ischaemia,
 CC allergic rhinitis, respiratory distress syndrome, endotoxin shock
 CC syndrome and atherosclerosis. The present sequence is the human alpha1
 CC integrin domain
 XX
 SQ Sequence 214 AA;

Query Match 1.3%; Score 15; DB 4; Length 214;
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
 |||||
 DB 22 DIVIVLDGNSIYPW 36

RESULT 59
 AAB53539
 ID AAB53539 standard; protein; 214 AA.
 XX
 AC AAB53539;
 DT 02-APR-2003 (first entry)
 XX
 DE Rat alpha 1-I domain protein #2.
 XX

Rat; very late activation antigen; VLA-1; beta1 containing integrin;
 KW immunological disorder; inflammatory disorder; skin related condition;
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;
 KW atherosclerosis; thyroiditis; aplastic anaemia; periarthritis nodosa;
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;
 KW graft versus host disease; myocardial ischaemia; alpha 1-I protein.
 XX
 OS Rattus sp.
 XX
 PN WO200283854-A2.
 XX
 PD 24-OCT-2002.
 XX

PF 12-APR-2002; 2002WO-US011521.
 XX
 PR 13-APR-2001; 2001US-0283794P.
 PR 06-JUL-2001; 2001US-0303689P.
 XX
 PA (BIOJ) BIOGEN INC.
 XX
 PI Lyne PD, Garber EA, Saldanha JW, Karpusas M;
 XX
 XX WPI; 2003-093009/08.
 DR

PT New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or
 PT fibrosis.
 XX
 PS Example 18; Fig 11A; 248pp; English.

XX The present invention relates to novel antibodies that specifically bind
 CC to very late activation (VLA-1; beta1 containing integrins) antigens and
 CC methods of using these antibodies to treat immunological disorders. The
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 CC mediated immunological or inflammatory disorders such as skin related
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal
 CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),
 CC periaarthritis nodosa, Hodgkin's disease, rheumatoid fever, osteoarthritis,
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or
 CC immediate hypersensitivity), graft and transplant rejections, graft
 CC versus host disease, conjunctivitis, swelling occurring after injury,
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is
 CC rat 1-I domain protein. This sequence is used in the exemplification of
 CC the invention

XX Sequence 214 AA;

Query Match 1.3%; Score 15; DB 6; Length 214;
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178
 |||||
 Db 22 DIVVLGSGNSIYPW 36

RESULT 60
 AAEE33540
 ID AAEE33540 standard; protein; 214 AA.

XX AAEE33540;
 DT 02-APR-2003 (first entry)
 XX Human alpha 1-I domain protein #2.
 XX Human; very late activation antigen; VLA-1; beta1 containing integrin;
 KW immunological disorder; inflammatory disorder; skin related condition;
 KW psoriasis; eczema; burn; dermatitis; respiratory distress syndrome;
 KW fibrosis; allergic rhinitis; asthma; bronchitis; tendonitis; bursitis;
 KW fever; migraine headache; inflammatory bowel disease; Crohn's disease;
 KW irritable bowel syndrome; colitis; colorectal cancer; vascular disease;
 KW atherosclerosis; thyroiditis; aplastic anaemia; periaarthritis nodosa;
 KW gastritis; Hodgkin's disease; rheumatic fever; autoimmune disease;
 KW osteoarthritis; type I diabetes; myasthenia gravis; rheumatoid arthritis;
 KW systemic lupus erythematosus; multiple sclerosis; nephrotic syndrome;
 KW renal failure; sarcoidosis; Behcet's syndrome; gingivitis; polymyositis;
 KW hypersensitivity; graft rejection; transplant rejection; conjunctivitis;
 KW graft versus host disease; myocardial ischaemia; alpha 1-I protein.

XX Homo sapiens.
 OS WO200283854-A2.
 XX 24-OCT-2002.
 XX 12-APR-2002; 2002WO-US011521.
 XX 13-APR-2001; 2001US-0283794P.
 XX 06-JUL-2001; 2001US-0303689P.
 XX (BIOJ) BIOGEN INC.
 XX Lyne PD, Garber EA, Saldanha JW, Karpusas M;

XX WPI; 2003-093009/08.
 XX New anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 PT mediated immunological or inflammatory disorders, e.g. psoriasis, eczema,
 PT burns, dermatitis, and abnormal proliferation of hair follicle cells or
 PT fibrosis.
 XX Example 18; Fig 12; 248pp; English.
 XX The present invention relates to novel antibodies that specifically bind
 CC to very late activation (VLA-1; beta1 containing integrins) antigens and
 CC methods of using these antibodies to treat immunological disorders. The
 CC anti-VLA-1 antibodies are useful for preventing or treating VLA-1-
 CC mediated immunological or inflammatory disorders such as skin related
 CC conditions (e.g. psoriasis, eczema, burns, dermatitis and abnormal
 CC proliferation of hair follicle cells), fibrosis (e.g. kidney or lung
 CC fibrosis), allergic rhinitis, respiratory distress syndrome, asthma,
 CC bronchitis, tendonitis, bursitis, fever, migraine headaches, gastro-
 CC intestinal conditions (e.g. inflammatory bowel disease, Crohn's disease,
 CC gastritis, irritable bowel syndrome, colitis and colorectal cancer),
 CC vascular diseases (e.g. atherosclerosis), thyroiditis, aplastic anaemia,
 CC periaarthritis nodosa, Hodgkin's disease, rheumatic fever, osteoarthritis,
 CC autoimmune diseases (e.g. type I diabetes, myasthenia gravis, rheumatoid
 CC arthritis, systemic lupus erythematosus and multiple sclerosis), renal
 CC failure, sarcoidosis, nephrotic syndrome, Behcet's syndrome, gingivitis,
 CC polymyositis, hypersensitivity (e.g. delayed type hypersensitivity or
 CC immediate hypersensitivity), graft and transplant rejections, graft
 CC versus host disease, conjunctivitis, swelling occurring after injury,
 CC myocardial ischaemia or endotoxin shock syndrome. The present sequence is
 CC human 1-I domain protein. This sequence is used in the exemplification of
 CC the invention

XX Sequence 214 AA;

Query Match 1.3%; Score 15; DB 6; Length 214;
 Best Local Similarity 100.0%; Pred. No. 9.5e-06;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178
 |||||
 Db 22 DIVVLGSGNSIYPW 36

RESULT 61
 ADEB6652
 ID ADEB6652 standard; protein; 1151 AA.

XX ADEB6652;
 XX 29-JAN-2004 (first entry)
 XX Human integrin alpha 1 subunit.
 DE human; secreted protein; cancer; liver disorder; hepatitis;
 KW neural disorder; Alzheimer's disease; integrin alpha 1 subunit.
 XX Homo sapiens.
 OS US2003129685-A1.
 XX 10-JUL-2003.
 XX 18-APR-2001; 2001US-00836353.
 XX 28-OCT-1998; 98US-0105971P.
 XX 27-OCT-1999; 99WO-US025031.
 XX 19-APR-2000; 2000US-0198407P.
 XX (NIJ) NI J.
 XX (YOUNG) YOUNG P E.
 XX (KENN) KENNY J J.
 XX (OLSE) OLSEN H S.

PA (MOORE/) MOORE P A.
PA (WEIY/) WEI Y.
PA (GREENE/) GREENE J M.
PA (RUBEN/) RUBEN S M.
XX Ni J, Young PE, Kenny JJ, Olsen HS, Moore PA, Wei Y, Greene JM;
PI Ruben SM;
XX WPI; 2004-020335/02.
XX New nucleic acid molecule, useful for preparing a medicament for
PT preventing, treating or ameliorating a medical condition e.g. cancer,
PT liver disorders or neural disorders.
XX PS Disclosure; SEQ ID NO 103; 380pp; English.
XX CC The invention relates to an isolated nucleic acid sequence, or its
CC allelic variant, a fragment of the cDNA sequence, or its fragment,
CC domain, epitope or species homologue. The nucleic acid is useful for
CC preparing a medicament for preventing, treating or ameliorating a medical
CC condition e.g., cancer, liver disorders such as hepatitis or neural
CC disorders such as Alzheimer's disease. The present sequence represents
CC the amino acid sequence of an integrin alpha 1 subunit.
XX SQ Sequence 1151 AA;
SQ
Query Match 1.3%; Score 15; DB 8; Length 1151;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 DIVIVLDGNSIYVP 178
DB 144 DIVIVLDGNSIYVP 158
RESULT 62
ABB90759
ID ABB90759 standard; protein; 1179 AA.
XX AC ABB90759;
XX 30-MAY-2002 (first entry)
XX Human Tumour Endothelial Marker polypeptide SEQ ID NO 250.
XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
XX normal endothelial marker; pan-endothelial marker; immunostimulant;
XX antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
XX polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
XX psoriasis.
XX Homo sapiens.
XX WO200210217-A2.
XX 07-FEB-2002.
XX 01-AUG-2001; 2001WO-US024031.
XX 02-AUG-2000; 2000US-0222599P.
XX 11-AUG-2000; 2000US-0224360P.
XX 11-APR-2001; 2001US-0282850P.
XX (UYJO) UNIV JOHNS HOPKINS.
XX St Croix B, Kinzler KW, Vogelstein B;
PI WPI; 2002-291856/33.
XX N-PSDB; ABL92113.
XX An isolated molecule comprising an antibody variable region which
PT specifically binds to an extracellular domain of a tumor endothelial
PT marker (TEM) protein, useful for inhibiting tumor growth.

XX Claim 35; Page 242-245; 331pp; English.
XX The invention relates to an isolated molecule comprising an antibody
CC variable region which specifically binds to an extracellular domain of a
CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,
CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
CC proteins have cytostatic, immunostimulant and antiangiogenic activity.
CC They are useful for inhibiting tumour growth, neoangiogenesis in subjects
CC bearing a vascularised tumour, polycystic kidney disease, diabetic
CC reuropathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM
CC genes and the encoded proteins (ABL92075-ABL92141 and ABB90721-ABB90789)
CC are disclosed, as are marker oligonucleotide sequences: tumour
CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal
CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers
CC (PEM) ABL91903-ABL91995
XX SQ Sequence 1179 AA;
SQ
Query Match 1.3%; Score 15; DB 5; Length 1179;
Best Local Similarity 100.0%; Pred. No. 4.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 DIVIVLDGNSIYVP 178
DB 172 DIVIVLDGNSIYVP 186
RESULT 63
ABP64915
ID ABP64915 standard; protein; 1179 AA.
XX AC ABP64915;
XX 25-FEB-2003 (first entry)
XX Human protein SEQ ID 575.
XX Human; expressed sequence tag; EST; haematopoietic disorder;
XX central nervous system disease; viral infection;
XX peripheral nervous system disease; non-healing wound; infectious disease;
XX immune deficiency; immune disorder; bacterial infection; allergy; cancer;
XX fungal infection; autoimmune disorder; coagulation disorder; neotropic;
XX antiallergic; antiinflammatory; immunosuppressive; neuroprotective;
XX cytostatic; haemostatic; virucide; antibacterial; fungicide;
XX immunostimulant; cerebroprotective.
XX Homo sapiens.
XX WO200259260-A2.
XX 01-AUG-2002.
XX 16-NOV-2001; 2001WO-US042950..
XX 17-NOV-2000; 2000US-00714936.
XX (HYSE-) HYSEQ INC.
XX Tang YT, Goodrich RW, Liu C, Zhou P, Asundi V, Zhang J, Zhao QA;
PI Ren F, Xue RJ, Yang Y, Wehrman T, Drmanac RT;
XX WPI; 2002-590824/63.
XX N-PSDB; ABQ99501.
XX New isolated polynucleotide, useful in research, diagnostic or
PT therapeutic methods, e.g. preventing or treating disorders involving
PT aberrant protein expression or biological activity.
XX Claim 20; SEQ ID NO 575; 394pp; English.
XX The present invention relates to novel human coding sequences (ABQ99269-
CC ABQ99608) and proteins (ABP64682-ABP65022). The sequences are useful in
CC

CC therapeutic, diagnostic and research methods. The polynucleotides may be
 CC used in the field of molecular biology as hybridisation probes, primers
 CC for PCR, for chromosome and gene mapping, for the recombinant production
 CC of protein, or in generation of anti-sense DNA or RNA. The
 CC polynucleotides are useful in diagnostics as expressed sequence tags
 CC (ESTs) for identifying expressed genes or for physical mapping of the
 CC human genome. The proteins may be used as molecular weight markers, or as
 CC nutritional sources or supplements. The proteins may be used to maintain
 CC and expand cell population in a totipotent or pluripotent state
 CC useful for re-engineering damaged or diseased tissues, transplantation,
 CC manufacture of bio-pharmaceuticals or the development of bio-sensors. The
 CC polynucleotides and proteins are useful for preventing, treating or
 CC ameliorating disorders involving aberrant protein expression or
 CC biological activity, e.g. haematopoietic disorders, central/peripheral
 CC nervous system diseases, mechanical and traumatic disorders, non-healing
 CC wounds, immune deficiencies and disorders, infectious diseases caused by
 CC viral, bacterial or fungal infection, autoimmune disorders, allergic
 CC reactions and conditions, coagulation disorders, or cancer. The
 CC polynucleotide sequences of the invention were assembled from ESTs
 CC isolated mainly by sequencing by hybridisation, and in some cases,
 CC sequences obtained from one or more public databases. Note: The sequence
 CC data for this patent did not form part of the printed specification, but
 CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 CC
 XX Sequence 1179 AA;

Query Match 1.3%; Score 15; DB 5; Length 1179;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178
 |||||
 Db 172 DIVVLGSGNSIYPW 186

RESULT 64
 ABUS4466
 ID ABUS4466 standard; protein; 1179 AA.

XX ABUS4466;

XX 12-MAR-2003 (first entry)

XX Human tumour endothelial marker TEM 30.

XX Human; endothelial cell; EC; tumour endothelial cell; TEM; NEM;
 XX Tumour endothelial marker; normal endothelial marker; PEM;
 XX Pan-endothelial marker; polycystic kidney disease; psoriasis;
 XX diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
 XX neovascularization; immune response; cytostatic; antidiabetic;
 XX ophthalmological; antirheumatic; antiarthritic; antipsoriatic.

XX Homo sapiens.

XX WO200283874-A2.

XX 24-OCT-2002.

XX 10-APR-2002; 2002WO-US008253.

XX 11-APR-2001; 2001US-0282850P.

XX 06-FEB-2002; 2002US-0354262P.

XX (UYJO) UNIV JOHNS HOPKINS.

XX Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;

XX WPI; 2003-093016/08.

XX N-PSDB; ABX72038.

XX New purified human transmembrane protein, designated as tumor endothelial
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors,
 PT

PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
 PT psoriasis.

XX Disclosure; Page 267-269; 374pp; English.

XX The present invention relates to a novel method for the isolation of
 CC endothelial cells (ECs), and the identification of genes expressed in
 CC normal and tumour ECs. Tumour endothelial marker (TEM), normal
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
 CC identified in human ECs. The human EC marker proteins and the
 CC polynucleotide sequences encoding them are useful for detecting,
 CC diagnosing or treating tumours as well as polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
 CC useful for inhibiting neoangiogenesis or tumour angiogenesis, for
 CC inducing an immune response to tumour endothelial cells in a patient, or
 CC for identifying candidate drugs for treating tumours. The present
 CC sequence represents a human TEM or NEM protein of the invention
 XX Sequence 1179 AA;

Query Match 1.3%; Score 15; DB 6; Length 1179;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVVLGSGNSIYPW 178
 |||||
 Db 172 DIVVLGSGNSIYPW 186

RESULT 65
 ABB90788
 ID ABB90788 standard; protein; 1180 AA.

XX ABB90788;

XX 30-MAY-2002 (first entry)

XX Rat Tumour Endothelial Marker polypeptide SEQ ID NO 307.

XX Human; mouse; rat; TEM; tumour endothelial marker; NEM; PEM; cytostatic;
 XX normal endothelial marker; pan-endothelial marker; immunostimulant;
 XX antiangiogenic; tumour; neoangiogenesis; vascularised tumour;
 XX polycystic kidney disease; diabetes; retinopathy; rheumatoid arthritis;
 XX psoriasis.

XX Rattus sp.

XX WO200210217-A2.

XX 07-FEB-2002.

XX 01-AUG-2001; 2001WO-US024031.

XX 02-AUG-2000; 2000US-0222599P.

XX 11-AUG-2000; 2000US-0224360P.

XX 11-APR-2001; 2001US-0282850P.

XX (UYJO) UNIV JOHNS HOPKINS.

XX St Croix B, Kinzler KW, Vogelstein B;

XX WPI; 2002-291856/33.

XX N-PSDB; ABL92141.

XX An isolated molecule comprising an antibody variable region which

PT specifically binds to an extracellular domain of a tumor endothelial
 PT marker (TEM) protein, useful for inhibiting tumor growth.

XX Disclosure; Page 314-318; 331pp; English.

XX The invention relates to an isolated molecule comprising an antibody
 CC variable region which specifically binds to an extracellular domain of a
 CC tumour endothelial marker (TEM) protein selected from ABB90732, ABB90740,

CC ABB90749, ABB90750 and ABB90769. The antibodies which bind to TEM
 CC proteins have cytostatic, immunostimulant and antiangiogenic activity.
 CC They are useful for inhibiting tumour growth, neoangiogenesis in subjects
 CC bearing a vascularised tumour, polycystic kidney disease, diabetic
 CC retinopathy, rheumatoid arthritis and psoriasis. Human, mouse and rat TEM
 CC genes and the encoded proteins (ABU92075-ABU92141 and ABB90721-ABB90789)
 CC are disclosed, as are marker oligonucleotide sequences: tumour
 CC endothelial markers (TEM) ABL91996-ABL92041 and ABL92143-ABL92191; normal
 CC endothelial markers (NEM) ABL92042-ABL92074; and pan-endothelial markers
 CC (PEM) ABL91903-ABL91995
 XX
 SQ Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 5; Length 1180;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 DIVIVLDGNSIYPW 178
 Db 172 DIVIVLDGNSIYPW 186
 |||||

RESULT 66
 ABU54495
 ID ABU54495 standard; protein; 1180 AA.
 XX
 AC ABU54495;

XX
 DT 12-MAR-2003 (first entry)
 XX
 DE Mouse tumour endothelial marker mTEM 30.
 XX
 KW Mouse; endothelial cell; EC; tumour endothelial cell; TEM; NEM;
 KW Tumour endothelial marker; normal endothelial marker; PEM;
 KW pan-endothelial marker; polycystic kidney disease; psoriasis;
 KW diabetic retinopathy; rheumatoid arthritis; tumour angiogenesis;
 KW neoangiogenesis; immune response; cytostatic; antidiabetic;
 KW ophthalmological; antirheumatic; antiarthritic; antipsoriatic.
 XX

OS Mus sp.

XX WO200283874-A2.
 XX
 XX 24-OCT-2002.
 XX
 XX 10-APR-2002; 2002WO-US008253.
 XX

PR 11-APR-2001; 2001US-0282850P.
 PR 06-FEB-2002; 2002US-0354262P.
 XX
 XX (UVJO) UNIV JOHNS HOPKINS.

XX Carson-Walter E, St Croix B, Kinzler KW, Vogelstein B;
 PI

XX WPI; 2003-093016/08.
 DR N-PSDB; ABX72066.

XX New purified human transmembrane protein, designated as tumor endothelial
 PT marker (TEM) 3, useful for detecting, diagnosing or treating tumors.
 PT polycystic kidney disease, diabetic retinopathy, rheumatoid arthritis or
 PT psoriasis.

XX Disclosure; Page 356-358; 374pp; English.

XX The present invention relates to a novel method for the isolation of
 CC endothelial cells (ECs), and the identification of genes expressed in
 CC normal and tumour ECs. Tumour endothelial marker (TEM) normal
 CC endothelial marker (NEM), and pan-endothelial marker (PEM) genes are
 CC identified in human ECs. The human EC marker proteins and the
 CC polynucleotide sequences encoding them are useful for detecting,
 CC diagnosing or treating tumours as well as polycystic kidney disease,
 CC diabetic retinopathy, rheumatoid arthritis, and psoriasis. They are also
 CC useful for inhibiting neoangiogenesis or tumour angiogenesis, for

CC inducing an immune response to tumour endothelial cells in a patient, or
 CC for identifying candidate drugs for treating tumours. The present
 CC sequence represents a mouse TEM protein

SQ Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 6; Length 1180;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 164 DIVIVLDGNSIYPW 178
 Db 172 DIVIVLDGNSIYPW 186
 |||||

RESULT 67
 ADE63568
 ID ADE63568 standard; protein; 1180 AA.

XX
 AC ADE63568;

XX
 DT 29-JAN-2004 (first entry)

XX Rat Protein P18614, SEQ ID NO 9512.

XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.

XX Rattus norvegicus.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.

XX (GEHO) GEN HOSPITAL CORP.

XX (FARB) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.
 DR GENBANK; P18614.

XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page: 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating

CC pain (e.g. spinal segmental nerve injury (Chung)), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of
 CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1180 AA;

Query Match 1.3%; Score 15; DB 7; Length 1180;
 Best Local Similarity 100.0%; Pred. No. 4.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGDSNSIYPW 178
 |||||
 Db 172 DIVVLGDSNSIYPW 186

RESULT 68
 AAB30928
 ID AAB30928 standard; peptide; 15 AA.

XX AC AAB30928;

XX DT 02-APR-2001 (first entry)

XX DE Antigenic peptide derived from a human alpha11 integrin chain.

XX KW Human; integrin; alpha11 subunit; fibroblast; muscle cell; chondrocyte;
 KW osteoblast; stem cell; cell damage; muscle dystrophy; fibrosis;
 KW wound healing; trauma; rheumatoid arthritis; osteoarthritis;
 KW osteoporosis; cartilage damage; bone damage; cartilage.

XX OS Homo sapiens.

XX PN WO200075187-A1.

XX PD 14-DEC-2000.

XX PF 31-MAY-2000; 2000WO-S0001135.

XX PR 03-JUN-1999; 99SE-00002056.

XX PA (ACTI-) ACTIVE BIOTECH AB.

XX PI Gullberg D;

XX DR WPI; 2001-071061/09.

XX PT Integrin subunit alpha 11 or integrin heterodimer comprising subunit
 PT alpha 11 in association with subunit beta, useful for treating muscle
 PT dystrophy, fibrosis, trauma, rheumatoid arthritis, and osteoarthritis.

XX PS Disclosure; Page 17; 79pp; English.

XX CC The present sequence is derived from the cytoplasmic domain of the human
 CC integrin subunit, designated alpha11. The alpha11 polynucleotide and
 CC polypeptide are useful as markers of cell target molecules, such as
 CC fibroblasts, muscle cells, chondrocytes, osteoblasts, mesenchymally
 CC derived cells or stem cells. They are also used for determining the
 CC differential-stage of cells during differentiation, development in
 CC pathological conditions, in tissue regeneration, in transplantation or in
 CC therapeutic and physiological repair of tissues. The pathological
 CC conditions involving subunit alpha11 are selected from damage of cells,
 CC muscle dystrophy, fibrosis, wound healing, trauma, rheumatoid arthritis,
 CC osteoarthritis and osteoporosis, damage of cartilage and bone, and
 CC cartilage and bone diseases. The polypeptide is useful for detecting the
 CC formation of cartilage during embryonic development, for detecting
 CC physiological therapeutic repair of cartilage and muscle, for selection
 CC and analysis, or for sorting, isolating or purification of chondrocytes
 CC and muscle cells, for detecting regeneration of cartilage or chondrocytes
 CC during transplantation of cartilage or chondrocytes during

CC transplantation of cartilage or chondrocytes, respectively, or of muscle
 CC or muscle cells during transplantation of muscle or muscle cells,
 CC respectively, and for studies of differentiation or chondrocytes or
 CC muscle cells

XX SQ Sequence 15 AA;

Query Match 1.2%; Score 14; DB 4; Length 15;
 Best Local Similarity 100.0%; Pred. No. 9.1e-06;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1175 RREFGLDPTPKVLE 1188
 |||||
 Db 2 RREFGLDPTPKVLE 15

RESULT 69

AAAY07728

ID AAAY07728 standard; protein; 1183 AA.

XX AC AAAY07728;

XX DT 01-JUL-1999 (first entry)

XX DE Armenian hamster alpha-1 integrin subunit protein.

XX KW VEGF; tumour angiogenesis inhibition; vascular endothelial growth factor;
 KW integrin cell surface receptor; capillary; blood vessel; hamster;
 KW alpha-1 subunit; alpha-2 subunit.

XX OS Cricetulus migratorius.

XX PN WO9916465-A1.

XX PD 08-APR-1999.

XX PF 30-SEP-1997; 97WO-US017485.

XX PR 30-SEP-1997; 97WO-US017485.

XX PA (BETH-) BETH ISRAEL DEACONESS MEDICAL CENT.

XX PA (SENG/) SENG D R.

XX PA (DETM/) DETMAR M.

XX PA (CLAF/) CLAFFEY K P.

XX PI Senger DR, Detmar M, Claffey KP;

XX DR WPI; 1999-254930/21.

XX PT Inhibition of tumor angiogenesis through interaction of vascular
 PT endothelial growth factor and integrin cell surface receptors.

XX PS Disclosure; Fig 2A-C; 64pp; English.

XX CC This invention describes a novel method for the inhibition of tumour
 CC angiogenesis mediated by vascular endothelial growth factor (VEGF) and
 CC integrin cell surface receptors expressed in vasculature of living
 CC subjects. The method inhibits new capillary and new blood vessel
 CC formation both within a tumour mass itself as well as in the immediately
 CC adjacent blood vasculature surrounding the perimeter of the tumour mass.
 CC interaction and dependence upon VEGF to induce specific integrin
 CC heterodimers in tumour angiogenesis provides a novel method for
 CC inhibiting tumour angiogenesis, and unlike prior art relies on the
 CC specific inter-relationship of VEGF and integrins, rather than
 CC concentrating solely on one specific class of protein

XX SQ Sequence 1183 AA;

Query Match 1.2%; Score 14; DB 2; Length 1183;
 Best Local Similarity 100.0%; Pred. No. 0.00042;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 164 DIVVLGDSNSIYP 177

AAY32243
 ID AAY32243 standard; protein; 1132 AA.
 AC AAY32243;
 XX 15-FEB-2000 (first entry)
 DT DT
 DE Human integrin subunit alpha-10 splice variant.
 XX Integrin alpha-10; ISa10; human; trauma; rheumatoid arthritis;
 XX osteoarthritis; osteoarthritis; cancer; atherosclerosis; inflammation;
 XX therapy; cartilage; chondrocyte; osteoblast; fibroblast; vaccine; marker;
 XX splice variant.
 XX Homo sapiens.
 OS
 PH Key Location/Qualifiers
 FT Peptide 1..22
 FT /note= "signal peptide"
 FT Protein 23..1132
 FT /note= "mature protein"
 XX WO9951639-A1.
 PN 14-OCT-1999.
 PD
 XX 31-MAR-1999; 99WO-SR000544.
 XX 02-APR-1998; 98SE-00001164.
 PR 28-JAN-1999; 99SE-00000319.
 XX (ACTI-) ACTIVE BIOTECH AB.
 PA
 XX Lundgren-Akerlund E;
 PI WPI; 2000-052639/04.
 DR N-PSDB; AAZ34720.
 XX New isolated integrin subunit alpha-10, used as a marker or target
 PT molecule for cells during development, regeneration and pathological
 PT conditions, e.g. arthritis, osteoarthritis, cancer, atherosclerosis or
 PT inflammation.
 PS Claim 1; Page 43-48; 90pp; English.
 XX This sequence represents a splice variant of novel human chondrocyte
 CC integrin subunit alpha-10 (ISa10). It is identical to ISa10 (see
 CC AAY32242) except for deletion of amino acids 975-986. The invention
 CC relates to a recombinant or isolated integrin heterodimer comprising the
 CC alpha10 subunit in association with subunit beta (especially beta-1). The
 CC heterodimer, subunit alpha-10 or splice variant can be used as a marker
 CC or target of all types of cells, e.g. of chondrocytes, osteoblasts and
 CC fibroblasts. They can also be used for treating pathological conditions
 CC involving ISa10, such as damage to cartilage, trauma, rheumatoid
 CC arthritis or osteoarthritis; for detecting the formation of cartilage
 CC during embryonal development, physiological or therapeutic repair of
 CC cartilage, or detecting regeneration of cartilage or chondrocytes during
 CC transplantation of cartilage or chondrocytes; for selection and analysis
 CC or for sorting, isolating or purification of chondrocytes and for in
 CC vitro studies of differentiation of chondrocytes; and as a target for
 CC anti-adhesive drugs or molecules in tendon, ligament, skeletal muscle or
 CC other tissues where adhesion impairs the function of the tissue (all
 CC claimed). ISa10 binding entities can be used to determine the
 CC differentiation-state of cells during embryonic development, angiogenesis
 CC or development of cancer, in pathological conditions such as rheumatoid
 CC arthritis, osteoarthritis or cancer, in tissue regeneration or in
 CC therapeutic and physiological repair of cartilage (claimed). A
 CC vaccine comprising the integrin heterodimer or subunit alpha-10 is also
 CC claimed. ISa10 polynucleotides, vectors, host cells and methods of
 CC producing recombinant ISa10 are also claimed
 XX Sequence 1132 AA;
 SQ
 Query Match 1.1%; Score 13; DB 3; Length 1132;
 Best Local Similarity 100.0%; Pred. No. 0.0039; Mismatches 0; Gaps 0;
 Matches 13; Conservative 0; Indels 0; Gaps 0;
 QY 166 VIVLDGNSIYPW 178
 DB 169 VIVLDGNSIYPW 181
 RESULT 73
 AAB64657
 ID AAB64657 standard; protein; 1152 AA.
 XX AAB64657;
 AC
 XX 22-MAR-2001 (first entry)
 DT Human secreted protein BLAST search protein SEQ ID NO: 167.
 DE
 XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
 KW vulnerary; anticonvulsant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.
 XX Homo sapiens.
 OS
 XX WO200077197-A1.
 PN 21-DEC-2000.
 PD
 XX 01-JUN-2000; 2000WO-US014934.
 PF 11-JUN-1999; 99US-0138599P.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA (ROSE/) ROSEN C A.
 XX Rosen CA, Ruben SM, Komatsoulis GA;
 PI WPI; 2001-032312/04.
 DR
 PT Isolated nucleic acid molecule encoding a human secreted protein is used
 PT in preventing, treating or ameliorating a medical condition.
 PS Disclosure; Page 543-546; 589pp; English.
 XX The invention relates to the isolation of genes AAF32757-F32803 encoding
 CC the human secreted proteins AAB64549-B64594. The sequence is a search
 CC result from a BLASTX homology search. The genes and proteins are useful
 CC for preventing, ameliorating or treating medical conditions, e.g. by
 CC protein or gene therapy. The genes are isolated from a range of human
 CC tissues disclosed in the specification. The nucleic acids, proteins,
 CC antibodies and (ant)agonists are useful in the diagnosis, treatment and
 CC prevention of: (a) cancer, e.g. breast and ovarian cancer, and other
 CC cancers of the adrenal gland, bone, bone marrow, breast, gastrointestinal
 CC tract, liver, lung, or urogenital; (b) immune disorders e.g. Addison's
 CC disease, allergies, autoimmune haemolytic anaemia, autoimmune
 CC thyroiditis, diabetes mellitus, Crohn's disease, multiple sclerosis,
 CC rheumatoid arthritis and ulcerative colitis; (c) cardiovascular disorders
 CC such as myocardial ischaemia; (d) wound healing; (e) neurological
 CC diseases e.g. cerebral anoxia and epilepsy; and (f) infectious diseases
 CC such as viral, bacterial, fungal and parasitic infections
 XX Sequence 1152 AA;
 SQ
 Query Match 1.1%; Score 13; DB 4; Length 1152;
 Best Local Similarity 100.0%; Pred. No. 0.004; Mismatches 0; Gaps 0;
 Matches 13; Conservative 0; Indels 0; Gaps 0;
 QY 166 VIVLDGNSIYPW 178
 DB 169 VIVLDGNSIYPW 181

RESULT 74

AAB64658
ID AAB64658 standard; protein; 1152 AA.
XX
AC AAB64658;
XX
DT 22-MAR-2001 (first entry)
XX
DE Human secreted protein BLAST search protein SEQ ID NO: 168.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; antitumor;
XX vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
FN WC2000077197-A1.
XX
PD 21-DEC-2000.
XX
PF 01-JUN-2000; 2000WO-US014934.
XX
PR 11-JUN-1999; 99US-0138599P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA (ROSE/) ROSEN C A.
XX
XX Rosen CA, Ruben SM, Komatsoulis GA;
PI
XX WPI; 2001-032312/04.
XX
XX Isolated nucleic acid molecule encoding a human secreted protein is used
XX in preventing, treating or ameliorating a medical condition.
XX
XX Disclosure; Page 547-551; 558pp; English.
XX
XX The invention relates to the isolation of genes AAF32757-F32803 encoding
XX the human secreted proteins AAB6458-B6459. The sequence is used as a
XX query sequence for doing BLASTX searches to identify homologous
XX sequences. The genes and proteins are useful for preventing, ameliorating
XX or treating medical conditions, e.g. by protein or gene therapy. The
XX genes are isolated from a range of human tissues disclosed in the
XX specification. The nucleic acids, proteins, antibodies and (antagonists
XX are useful in the diagnosis, treatment and prevention of: (a) cancer,
XX e.g. breast and ovarian cancer, and other cancers of the adrenal gland,
XX bone, bone marrow, breast, gastrointestinal tract, liver, lung, or
XX urogenital; (b) immune disorders e.g. Addison's disease, allergies,
XX autoimmune haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
XX Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
XX colitis; (c) cardiovascular disorders such as myocardial ischaemias; (d)
XX wound healing; (e) neurological diseases e.g. cerebral anoxia and
XX epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
XX and parasitic infections
XX
SQ Sequence 1152 AA;

Query Match 1.1%; Score 13; DB 4; Length 1152;
Best Local Similarity 100.0%; Pred. No. 0.004;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSIYPW 178
|||
Db 169 VIVLDGNSIYPW 181

RESULT 75
AAY32242
ID AAY32242 standard; protein; 1167 AA.
XX

AAY32242;
XX
XX 15-FEB-2000 (first entry)
DE Human integrin subunit alpha-10.
XX
XX Integrin alpha-10; Isal0; human; trauma; rheumatoid arthritis;
XX osteoarthritis; osteoarthritis; cancer; atherosclerosis; inflammation;
XX therapy; cartilage; chondrocyte; osteoblast; fibroblast; vaccine; marker.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..22
FT /note= "signal peptide"
FT Protein 23..1145
FT /note= "mature protein"
FT Domain 23..1120
FT /note= "extracellular domain"
FT Modified-site 98
FT /note= "N-glycosylated"
FT Domain 162..359
FT /note= "I-domain"
FT Modified-site 336
FT /note= "N-glycosylated"
FT Modified-site 364
FT /note= "N-glycosylated"
FT Binding-site 494..502
FT /note= "cation binding site motif"
FT Binding-site 558..566
FT /note= "cation binding site motif"
FT Binding-site 620..628
FT /note= "cation binding site motif"
FT Modified-site 733
FT /note= "N-glycosylated"
FT Modified-site 839
FT /note= "N-glycosylated"
FT Modified-site 921
FT /note= "N-glycosylated"
FT Modified-site 1018
FT /note= "N-glycosylated"
FT Modified-site 1039
FT /note= "N-glycosylated"
FT Domain 1121..1145
FT /note= "transmembrane domain"
FT Domain 1122..1167
FT /note= "cytoplasmic domain, specifically claimed in Claim
21"

WO9951639-A1.
14-OCT-1999.
31-MAR-1999; 99WO-SE000544.
02-APR-1998; 98SE-00001164.
28-JAN-1999; 99SE-00000319.
(ACTI-) ACTIVE BIOTECH AB.
Lundgren-Akerlund E;
WPI; 2000-052639/04.
N-PSDB; AA234719.
New isolated integrin subunit alpha-10, used as a marker or target
molecule for cells during development, regeneration and pathological
conditions, e.g. arthritis, osteoarthritis, cancer, atherosclerosis or
inflammation.
Claim 1; Fig 6; 90pp; English.
This sequence represents novel human chondrocyte integrin subunit alpha-

CC 10 (ISa10). A splice variant is given in AAY32243. The invention relates
CC to a recombinant or isolated integrin heterodimer comprising the alpha10
CC subunit in association with subunit beta (especially beta-1). The
CC heterodimer and the subunit alpha-10 can be used as markers or targets of
CC all types of cells, e.g. of chondrocytes, osteoblasts and fibroblasts.
CC They can also be used for treating pathological conditions involving
CC ISa10, such as damage to cartilage, trauma, rheumatoid arthritis or
CC osteoarthritis; for detecting the formation of cartilage during embryonal
CC development, physiological or therapeutic repair of cartilage, or
CC detecting regeneration of cartilage or chondrocytes during
CC transplantation of cartilage or chondrocytes; for selection and analysis
CC or for sorting, isolating or purification of chondrocytes and for in
CC vitro studies of differentiation of chondrocytes; and as a target for
CC anti-adhesive drugs or molecules in tendon, ligament, skeletal muscle or
CC other tissues where adhesion impairs the function of the tissue (all
CC claimed). ISa10 binding entities can be used to determine the
CC differentiation-state of cells during embryonic development, angiogenesis
CC or development of cancer, in pathological conditions such as rheumatoid
CC arthritis, osteoarthritis or cancer, in tissue regeneration or in
CC therapeutic and physiological repair of cartilage (claimed). A
CC vaccine comprising the integrin heterodimer or subunit alpha-10 is also
CC claimed. ISa10 polynucleotides, vectors, host cells and methods of
CC producing recombinant ISa10 are also claimed

XX
SQ Sequence 1167 AA;

Query Match 1.1%; Score 13; DB 3; Length 1167;
Best Local Similarity 100.0%; Pred. No. 0.0041;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 166 VIVLDGNSIYPW 178
Db 169 VIVLDGNSIYPW 181
|||||

Search completed: September 21, 2004, 13:31:38
Job time : 156 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:42:08 ; Search time 79 Seconds

(without alignments)
4829.194 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRGLVVAWSLWPGFT.....FRSARRRRPGLDPTPKVLE 1188

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1342398 seqs, 321133274 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1342398

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database :

Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	926	77.9	1188	15	US-10-291-265-338
2	911	76.7	1188	15	US-10-291-265-810
3	911	76.7	1189	10	US-09-984-130-35
4	911	76.7	1189	10	US-09-836-353A-35
5	911	76.7	1189	12	US-10-262-839-4
6	810	68.2	1034	10	US-09-984-130-43
7	810	68.2	1034	10	US-09-836-353A-43
8	753	63.4	1120	12	US-10-262-839-6
9	489	41.2	707	9	US-09-764-870-313
10	489	41.2	707	14	US-10-125-540-313
11	245	20.6	437	15	US-10-108-260A-3386
12	193	16.2	193	10	US-09-805-354-8
13	193	16.2	193	12	US-09-758-493-8
14	193	16.2	193	14	US-10-144-259-8
15	88	7.4	103	9	US-09-764-870-472

16	88	7.4	103	11	US-09-764-875-1193	Sequence 1193, Ap
17	88	7.4	103	14	US-10-125-540-472	Sequence 472, App
18	42	3.5	42	14	US-10-144-259-25	Sequence 25, Appl
19	42	3.5	545	10	US-09-866-050A-500	Sequence 500, App
20	42	3.5	688	10	US-09-866-050A-524	Sequence 624, App
21	42	3.5	696	10	US-09-866-050A-501	Sequence 501, App
22	33	2.8	33	10	US-09-984-130-102	Sequence 102, App
23	33	2.8	33	10	US-09-836-353A-102	Sequence 102, App
24	21	1.8	158	10	US-09-866-050A-391	Sequence 391, App
25	17	1.4	85	10	US-09-866-050A-390	Sequence 390, App
26	15	1.3	148	9	US-09-764-870-284	Sequence 284, App
27	15	1.3	148	9	US-09-764-870-444	Sequence 444, App
28	15	1.3	148	14	US-10-125-540-284	Sequence 284, App
29	15	1.3	148	14	US-10-125-540-444	Sequence 444, App
30	15	1.3	132	12	US-10-346-863-59	Sequence 59, Appl
31	15	1.3	132	16	US-10-474-832-57	Sequence 57, Appl
32	15	1.3	132	16	US-10-474-832-60	Sequence 60, Appl
33	15	1.3	132	16	US-10-474-832-61	Sequence 61, Appl
34	15	1.3	132	16	US-09-805-354-5	Sequence 5, Appl
35	15	1.3	132	16	US-09-758-493-5	Sequence 5, Appl
36	15	1.3	132	16	US-10-144-259-5	Sequence 5, Appl
37	15	1.3	212	9	US-09-996-738-5	Sequence 5, Appl
38	15	1.3	212	9	US-09-996-738-6	Sequence 6, Appl
39	15	1.3	214	12	US-10-625-260-5	Sequence 5, Appl
40	15	1.3	214	12	US-10-625-260-6	Sequence 6, Appl
41	15	1.3	214	12	US-10-625-260-9	Sequence 9, Appl
42	15	1.3	214	13	US-10-061-658-5	Sequence 5, Appl
43	15	1.3	214	13	US-10-061-658-6	Sequence 6, Appl
44	15	1.3	214	13	US-10-061-658-9	Sequence 9, Appl
45	15	1.3	214	16	US-10-474-832-63	Sequence 63, Appl
46	15	1.3	214	16	US-10-474-832-64	Sequence 64, Appl
47	15	1.3	1151	10	US-09-984-130-103	Sequence 103, App
48	15	1.3	1151	10	US-09-836-353A-103	Sequence 103, App
49	15	1.3	1119	12	US-09-918-715-250	Sequence 250, App
50	15	1.3	1180	12	US-09-918-715-307	Sequence 307, App
51	13	1.1	195	10	US-09-805-354-7	Sequence 7, Appl
52	13	1.1	195	12	US-09-758-493-7	Sequence 7, Appl
53	13	1.1	1167	16	US-10-144-259-7	Sequence 7, Appl
54	13	1.1	1167	16	US-10-144-259-9	Sequence 9, Appl
55	13	1.1	1177	16	US-10-741-601-531	Sequence 531, App
56	11	0.9	43	14	US-10-741-601-532	Sequence 532, App
57	11	0.9	1147	12	US-10-336-603A-42	Sequence 42, Appl
58	11	0.9	1181	12	US-10-211-462-187	Sequence 187, App
59	11	0.9	1181	14	US-10-160-354-2	Sequence 2, Appl
60	11	0.9	1181	15	US-10-295-027-1286	Sequence 1286, Ap
61	9	0.8	164	15	US-10-108-260A-3415	Sequence 3415, Ap
62	9	0.8	312	16	US-10-038-854-102	Sequence 102, App
63	9	0.8	312	16	US-10-369-493-2470	Sequence 2470, Ap
64	8	0.7	17	9	US-09-350-259-20	Sequence 20, Appl
65	8	0.7	17	10	US-09-891-943-20	Sequence 20, Appl
66	8	0.7	43	14	US-10-144-259-27	Sequence 27, Appl
67	8	0.7	67	9	US-09-978-295A-85	Sequence 85, Appl
68	8	0.7	67	9	US-09-978-697-85	Sequence 85, Appl
69	8	0.7	67	9	US-09-978-192A-85	Sequence 85, Appl
70	8	0.7	67	9	US-09-999-832A-85	Sequence 85, Appl
71	8	0.7	67	10	US-09-978-189-85	Sequence 85, Appl
72	8	0.7	67	10	US-09-978-608A-85	Sequence 85, Appl
73	8	0.7	67	10	US-09-978-585A-85	Sequence 85, Appl
74	8	0.7	67	10	US-09-978-191A-85	Sequence 85, Appl
75	8	0.7	67	10	US-09-978-403A-85	Sequence 85, Appl
76	8	0.7	67	10	US-09-978-564A-85	Sequence 85, Appl
77	8	0.7	67	10	US-09-999-833A-85	Sequence 85, Appl
78	8	0.7	67	10	US-09-981-915A-85	Sequence 85, Appl
79	8	0.7	67	10	US-09-978-824-85	Sequence 85, Appl
80	8	0.7	67	10	US-09-918-585A-85	Sequence 85, Appl
81	8	0.7	67	10	US-09-978-423A-85	Sequence 85, Appl
82	8	0.7	67	10	US-09-978-193A-85	Sequence 85, Appl
83	8	0.7	67	10	US-09-999-830A-85	Sequence 85, Appl
84	8	0.7	67	10	US-09-978-757A-85	Sequence 85, Appl
85	8	0.7	67	10	US-09-978-187B-85	Sequence 85, Appl
86	8	0.7	67	10	US-09-978-643A-85	Sequence 85, Appl
87	8	0.7	67	10	US-09-978-375A-85	Sequence 85, Appl
88	8	0.7	67	10	US-09-978-298A-85	Sequence 85, Appl

89 8 0.7 67 10 US-09-978-188A-85 Sequence 85, Appl
90 8 0.7 67 10 US-09-978-681A-85 Sequence 85, Appl
91 8 0.7 67 10 US-09-978-194A-85 Sequence 85, Appl
92 8 0.7 67 10 US-09-999-829A-85 Sequence 85, Appl
93 8 0.7 67 10 US-09-978-299A-85 Sequence 85, Appl
94 8 0.7 67 10 US-09-978-665A-85 Sequence 85, Appl
95 8 0.7 67 10 US-09-978-802A-85 Sequence 85, Appl
96 8 0.7 67 12 US-10-164-749A-85 Sequence 324, App
97 8 0.7 67 12 US-10-221-278-324 Sequence 85, Appl
98 8 0.7 67 12 US-09-993-831A-85 Sequence 85, Appl
99 8 0.7 67 12 US-10-013-917A-85 Sequence 85, Appl
100 8 0.7 67 12 US-10-013-917A-85 Sequence 85, Appl

ALIGNMENTS

RESULT 1
US-10-291-265-338
; Sequence 338, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; PRIOR FILING DATE: 2000-01-25
; PRIOR FILING DATE: 2000-01-25
; PRIOR FILING DATE: 2000-01-25
; PRIOR FILING DATE: 2000-01-25
; PRIOR FILING DATE: 2000-07-17
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-08-03
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 338
; LENGTH: 1188
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-291-265-338

Query Match 77.9%; Score 926; DB 15; Length 1188;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1126; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 MDLPRGLVAVALSMPGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLTVGA 60
DB 1 MDLPRGLVAVALSMPGFTDTFNMDTRKPRVPGSRTAFPGYTVQOHDISGNKWLTVGA 60
QY 61 PLETNGYQKTDVYKCPVTHGCTKLNIGRVTLNSVSEKDNMRGLSLATNPKNSELA 120
DB 61 PLETNGYQKTDVYKCPVTHGCTKLNIGRVTLNSVSEKDNMRGLSLATNPKNSELA 120
QY 121 CSPLSHCEGSSYITGTCMSRVNSFRSKTVAPALQRCQTYMDIVILVLDGNSIYPWVE 180
DB 121 CSPLSHCEGSSYITGTCMSRVNSFRSKTVAPALQRCQTYMDIVILVLDGNSIYPWVE 180
QY 181 VOHFLINILKYYIGPQIQVQVQGVGDVHFEHNLNDRSVKDVVEAASHIEQGGTET 240
DB 181 VOHFLINILKYYIGPQIQVQVQGVGDVHFEHNLNDRSVKDVVEAASHIEQGGTET 240
QY 241 RTAFGTEFARSAFQKGRKAKVMIVITDGEHSDSPLEKVIQOQSRDNTVYAVAVL 300
DB 241 RTAFGTEFARSAFQKGRKAKVMIVITDGEHSDSPLEKVIQOQSRDNTVYAVAVL 300
QY 301 GYNNRRGINPFTLNEIKYIASDPDDKHPFNVTDEAALKDVIDALGDRIFSLEGTNKNET 360
DB 301 GYNNRRGINPFTLNEIKYIASDPDDKHPFNVTDEAALKDVIDALGDRIFSLEGTNKNET 360
QY 361 SFGLEMSQTFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPEELKN 420

DB 361 SFGLEMSQTFSSHVVEDGVLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEPPEELKN 420
QY 421 HGAYLGVTVTSVSSRQGRVYVAGAPRFNHTKVVILFTVHNRSRTIHOAMGQOIGSYF 480
DB 421 HGAYLGVTVTSVSSRQGRVYVAGAPRFNHTKVVILFTVHNRSRTIHOAMGQOIGSYF 480
QY 481 GSBITSVDIDGCVTDVLLVAGPMYNEGRERKGVVYELRQNRQNFYNGTLKDSHSYQNA 540
DB 481 GSBITSVDIDGCVTDVLLVAGPMYNEGRERKGVVYELRQNRQNFYNGTLKDSHSYQNA 540
QY 541 RFSSIASVARDLNQDSYNDVVGAPLEDNHAIAIYIFHGFRCGSIILTKPKORTASBELATG 600
DB 541 RFSSIASVARDLNQDSYNDVVGAPLEDNHAIAIYIFHGFRCGSIILTKPKORTASBELATG 600
QY 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQVQINASLHFEPSKINI FHRDCK 660
DB 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQVQINASLHFEPSKINI FHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGIIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720
DB 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGIIRYNATMDERRYTTPRAHLDEGGDRFTNRAV 720
QY 721 LLSGGQELCERINFHVLDTADYVKPTFSVEYSLEDPDHGPMDDGWPTTLRVSVFPFNG 780
DB 721 LLSGGQELCERINFHVLDTADYVKPTFSVEYSLEDPDHGPMDDGWPTTLRVSVFPFNG 780
QY 781 CNEDEHCVPLVDLARSDLPTAMEYQCVLRPAQDCSAYTILSFTDTVPIIESTQRVAV 840
DB 781 CNEDEHCVPLVDLARSDLPTAMEYQCVLRPAQDCSAYTILSFTDTVPIIESTQRVAV 840
QY 841 EATLENRGENAYSTVLNISOANLQFASLIQKEDSDGSECVNEERRLOKQVCNVSYPF 900
DB 841 EATLENRGENAYSTVLNISOANLQFASLIQKEDSDGSECVNEERRLOKQVCNVSYPF 900
QY 901 RAKAKVAFRLDSEFSKSIPLHLEIELAAGSDSNRSDTKEDNVAFLRHLKYEADVLFT 960
DB 901 RAKAKVAFRLDSEFSKSIPLHLEIELAAGSDSNRSDTKEDNVAFLRHLKYEADVLFT 960
QY 961 RSSLSHYEVKPNSSLYRYDGGIPPPSCIFRQNLGLFIHGMKMTIPIATRSNRL 1020
DB 961 RSSLSHYEVKPNSSLYRYDGGIPPPSCIFRQNLGLFIHGMKMTIPIATRSNRL 1020
QY 1021 KLDRFLTDEANTSCNIWGNSTYRPTFVEEDLRRAPQLNHSNSDVVSINCLRLVNPQEI 1080
DB 1021 KLDRFLTDEANTSCNIWGNSTYRPTFVEEDLRRAPQLNHSNSDVVSINCLRLVNPQEI 1080
QY 1081 NFLLGNLWLSRLKALKYKSMKIMVNAALQROFHSPPFIREDPSRQI 1128
DB 1081 NFLLGNLWLSRLKALKYKSMKIMVNAALQROFHSPPFIREDPSRQI 1128

RESULT 2

US-10-291-265-810
; Sequence 810, Application US/10291265
; Publication No. US20030232054A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang et al
; TITLE OF INVENTION: No. US20030232054A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-017 (785)
; CURRENT APPLICATION NUMBER: US/10/291,265
; CURRENT FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/491,404
; PRIOR FILING DATE: 2000-01-25
; PRIOR APPLICATION NUMBER: 09/617,746
; PRIOR FILING DATE: 2000-07-17
; PRIOR APPLICATION NUMBER: 09/631,451
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: 09/633,870
; PRIOR FILING DATE: 2000-09-15
; NUMBER OF SEQ ID NOS: 944
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 810																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; LENGTH: 1188																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; TYPE: PRT																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
; ORGANISM: Homo sapiens																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
US-10-291-265-810																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Query Match																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Best Local Similarity 76.7%; Score 911; DB 15; Length 1188;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																														
QY	1	MDLPRGLVAVALS	LPFGTDTFNM	TRKPRVIPS	RTAFPGYT	VOQHDIS	GNKWL	VVGA	60																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
DB	1	MDLPRGLVAVALS	LPFGTDTFNM	TRKPRVIPS	RTAFPGYT	VOQHDIS	GNKWL	VVGA	60																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
QY	61	PLETNGYQKTGDVYK	CPVTHGNC	TKNLGRV	TLNSVSEK	DNMRGL	SLATNP	KDNSFLA	120																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
DB	61	PLETNGYQKTGDVYK	CPVTHGNC	TKNLGRV	TLNSVSEK	DNMRGL	SLATNP	KDNSFLA	120																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
QY	121	CSPLMSHECGSS	YTTGMC	SRVNSN	FRFSKT	VAPALQ	RCQTYMD	IVVL	DGNSIYPWVE	180																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
DB	121	CSPLMSHECGSS	YTTGMC	SRVNSN	FRFSKT	VAPALQ	RCQTYMD	IVVL	DGNSIYPWVE	180																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
QY	181	VQHFLINILK	KFYIGP	QIQGVVQ	YGEDV	VHBEFL	NDYRS	VKD	VVEAASHIE	QGGTET	240																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
DB	181	VQHFLINILK	KFYIGP	QIQGVVQ	YGEDV	VHBEFL	NDYRS	VKD	VVEAASHIE	QGGTET	240																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
QY	241	RTAFGIEFAR	SEAFQKGR	GKAKVM	IVITD	GESHDS	PDLEK	VI	QOSERD	NVTRYAV	300																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
DB	241	RTAFGIEFAR	SEAFQKGR	GKAKVM	IVITD	GESHDS	PDLEK	VI	QOSERD	NVTRYAV	300																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																			
QY	301	GYNRRGIN	PTFLNE	IKYIAS	DDPKH	FNVT	DEAAL	KD	IVDALG	DRIF	LEGTKNET	360																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
DB	301	GYNRRGIN	PTFLNE	IKYIAS	DDPKH	FNVT	DEAAL	KD	IVDALG	DRIF	LEGTKNET	360																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																		
QY	361	SFGLMSQ	TGFSSHV	VEDG	VLLG	AVGAYD	WNGAVL	KET	SAGK	VI	PLRESYL	KEPPEELKN	420																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
DB	361	SFGLMSQ	TGFSSHV	VEDG	VLLG	AVGAYD	WNGAVL	KET	SAGK	VI	PLRESYL	KEPPEELKN	420																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
QY	421	HGAYLG	YTVT	SVSSR	QGRVY	VAGAP	RFNHT	KGK	VI	FTMHN	NRSLT	THQMRG	QOIGSYF	480																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
DB	421	HGAYLG	YTVT	SVSSR	QGRVY	VAGAP	RFNHT	KGK	VI	FTMHN	NRSLT	THQMRG	QOIGSYF	480																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																
QY	481	GSET	TSVD	ID	DG	GYT	DVLL	VCAP	MYF	NEGR	ERG	KVYV	YEL	RQNF	RVNG	TLKDSH	SQNA	540																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
DB	481	GSET	TSVD	ID	DG	GYT	DVLL	VCAP	MYF	NEGR	ERG	KVYV	YEL	RQNF	RVNG	TLKDSH	SQNA	540																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
QY	541	RFSS	SIAS	VR	DLN	QDS	YND	VV	VGAP	LED	NHAG	AIYI	FH	GFRG	SILK	TPK	ORIT	ASE	LATG	600																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
DB	541	RFSS	SIAS	VR	DLN	QDS	YND	VV	VGAP	LED	NHAG	AIYI	FH	GFRG	SILK	TPK	ORIT	ASE	LATG	600																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
QY	601	LQYFGCS	IHGQ	DLN	EDG	LI	DLA	VAGAL	GN	AVL	WSPV	QV	IN	AS	LHP	FPSK	IN	I	FHRD	CK	660																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
DB	601	LQYFGCS	IHGQ	DLN	EDG	LI	DLA	VAGAL	GN	AVL	WSPV	QV	IN	AS	LHP	FPSK	IN	I	FHRD	CK	660																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
QY	661	RSRGRAT	CLAA	FL	CF	TP	IF	LPH	FO	TT	VG	IR	YN	AT	MD	ERR	Y	TP	RAH	L	DEG	DR	FT	NR	AV	720																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
DB	661	RSRGRAT	CLAA	FL	CF	TP	IF	LPH	FO	TT	VG	IR	YN	AT	MD	ERR	Y	TP	RAH	L	DEG	DR	FT	NR	AV	720																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
QY	721	LASSG	QEL	CER	IN	FH	VL	D	AD	Y	V	K	P	T	F	S	V	E	S	L	E	D	P	H	G	M	L	D	G	W	T	T	L	R	V	S	V	P	F	N	W	780																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
DB	721	LASSG	QEL	CER	IN	FH	VL	D	AD	Y	V	K	P	T	F	S	V	E	S	L	E	D	P	H	G	M	L	D	G	W	T	T	L	R	V	S	V	P	F	N	W	780																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																				
QY	781	CNEDEH	CV	DL	AD	RS	DL	PT	AME	Y	C	Q	R	V	L	R	PAQ	D	C	S	A	Y	T	L	S	F	O	T	T	F	I	E	S	T	R	Q	R	V	A	V	840																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
DB	781	CNEDEH	CV	DL	AD	RS	DL	PT	AME	Y	C	Q	R	V	L	R	PAQ	D	C	S	A	Y	T	L	S	F	O	T	T	F	I	E	S	T	R	Q	R	V	A	V	840																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					
QY	841	EATLEN	RG	EN	AY	S	T	V	L	N	I	S	Q	S	A	N	L	Q	P	A	S	L	I	Q	E	D	S	D	G	S	I	E	C	V	N	E	R	R	L	Q	K	V	C	N	V	S	Y	P	F	F	900																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
DB	841	EATLEN	RG	EN	AY	S	T	V	L	N	I	S	Q	S	A	N	L	Q	P	A	S	L	I	Q	E	D	S	D	G	S	I	E	C	V	N	E	R	R	L	Q	K	V	C	N	V	S	Y	P	F	F	900																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																											
QY	901	RAKAK	V	A	F	R	L	D	911																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																					

Query Match
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MDLPRGLVAVALS	LPFGTDTFNM	TRKPRVIPS	RTAFPGYT	VOQHDIS	GNKWL	VVGA	60											
DB	1	MDLPRGLVAVALS	LPFGTDTFNM	TRKPRVIPS	RTAFPGYT	VOQHDIS	GNKWL	VVGA	60											
QY	61	PLETNGYQKTGDVYK	CPVTHGNC	TKNLGRV	TLNSVSEK	DNMRGL	SLATNP	KDNSFLA	120											
DB	61	PLETNGYQKTGDVYK	CPVTHGNC	TKNLGRV	TLNSVSEK	DNMRGL	SLATNP	KDNSFLA	120											
QY	121	CSPLMSHECGSS	YTTGMC	SRVNSN	FRFSKT	VAPALQ	RCQTYMD	IVVL	DGNSIYPWVE	180										
DB	121	CSPLMSHECGSS	YTTGMC	SRVNSN	FRFSKT	VAPALQ	RCQTYMD	IVVL	DGNSIYPWVE	180										
QY	181	VQHFLINILK	KFYIGP	QIQGVVQ	YGEDV	VHBEFL	NDYRS	VKD	VVEAASHIE	QGGTET	240									
DB	181	VQHFLINILK	KFYIGP	QIQGVVQ	YGEDV	VHBEFL	NDYRS	VKD	VVEAASHIE	QGGTET	240									
QY	241	RTAFGIEFAR	SEAFQKGR	GKAKVM	IVITD	GESHDS	PDLEK	VI	QOSERD	NVTRYAV	300									
DB	241	RTAFGIEFAR	SEAFQKGR	GKAKVM	IVITD	GESHDS	PDLEK	VI	QOSERD	NVTRYAV	300									
QY	301	GYNRRGIN	PTFLNE	IKYIAS	DDPKH	FNVT	DEAAL	KD	IVDALG	DRIF	LEGTKNET	360								
DB	301	GYNRRGIN	PTFLNE	IKYIAS	DDPKH	FNVT	DEAAL	KD	IVDALG	DRIF	LEGTKNET	360								
QY	361	SFGLMSQ	TGFSSHV	VEDG	VLLG	AVGAYD	WNGAVL	KET	SAGK	VI	PLRESYL	KEPPEELKN	420							
DB	361	SFGLMSQ	TGFSSHV	VEDG	VLLG	AVGAYD	WNGAVL	KET	SAGK	VI	PLRESYL	KEPPEELKN	420							
QY	421	HGAYLG	YTVT	SVSSR	QGRVY	VAGAP	RFNHT	KGK	VI	FTMHN	NRSLT	THQMRG	QOIGSYF	480						
DB	421	HGAYLG	YTVT	SVSSR	QGRVY	VAGAP	RFNHT	KGK	VI	FTMHN	NRSLT	THQMRG	QOIGSYF	480						
QY	481	GSET	TSVD	ID	DG	GYT	DVLL	VCAP	MYF	NEGR	ERG	KVYV	YEL	RQNF	RVNG	TLKDSH	SQNA	540		
DB	481	GSET	TSVD	ID	DG	GYT	DVLL	VCAP	MYF	NEGR	ERG	KVYV	YEL	RQNF	RVNG	TLKDSH	SQNA	540		
QY	541	RFSS	SIAS	VR	DLN	QDS	YND	VV	VGAP	LED	NHAG	AIYI	FH	GFRG	SILK	TPK	ORIT	ASE	LATG	600
DB	541	RFSS	SIAS	VR	DLN	QDS	YND	VV	VGAP	LED	NHAG	AIYI	FH	GFRG	SILK	TPK	ORIT	ASE	LATG	600

QY 601 LQYFGCSIHGQDLNEDGLIDLAVALGNAVILWSRPVQINASLHFEPSKINIFHRDCK 660
Db 601 LQYFGCSIHGQDLNEDGLIDLAVALGNAVILWSRPVQINASLHFEPSKINIFHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIYNATMDERRVTPRAHLDEGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIYNATMDERRVTPRAHLDEGDRFTNRAV 720
QY 721 LLSGGQELCERINFHVLDADYVVPVTSFVSVEYSLSDPDHGPMLDDGWPPTTLRVSVYPPFWG 780
Db 721 LLSGGQELCERINFHVLDADYVVPVTSFVSVEYSLSDPDHGPMLDDGWPPTTLRVSVYPPFWG 780
QY 781 CNEDEHCVPDLVLDARSDDLPTAMEYCORVLRKPAQDCSAYTILSPDTTTFIESTRQAV 840
Db 781 CNEDEHCVPDLVLDARSDDLPTAMEYCORVLRKPAQDCSAYTILSPDTTTFIESTRQAV 840
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERRLQKQVCNVSYPFF 900
Db 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERRLQKQVCNVSYPFF 900
QY 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911

RESULT 4

US-09-836-353A-35
; Sequence 35, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P1
; CURRENT APPLICATION NUMBER: US/09/836,353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 35
; LENGTH: 1189
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-35

Query Match 76.7%; Score 911; DB 10; Length 1189;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 911; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MOLPRLVAVALSLSWPGTDFNWDTKPRVPGSRATFTCYTVOOHDISGNKWLAVGA 60
Db 1 MOLPRLVAVALSLSWPGTDFNWDTKPRVPGSRATFTCYTVOOHDISGNKWLAVGA 60
QY 61 PLETNQYQKTDGVYKCPVIGHNCTKLNLRVTLNVSERKONMRGLSLATNPKDNSFLA 120
Db 61 PLETNQYQKTDGVYKCPVIGHNCTKLNLRVTLNVSERKONMRGLSLATNPKDNSFLA 120
QY 121 CSPLASHEGCSYITTCMSRVNSFRSKTVAPALQRCQYTMDIVIVLDGNSIYPPWE 180
Db 121 CSPLASHEGCSYITTCMSRVNSFRSKTVAPALQRCQYTMDIVIVLDGNSIYPPWE 180
QY 181 VQHFILNLIKXYIGPGQIQVGVQYGEDVHVEFLNLYRSVKDVVEAASHIEQGGTET 240
Db 181 VQHFILNLIKXYIGPGQIQVGVQYGEDVHVEFLNLYRSVKDVVEAASHIEQGGTET 240
QY 241 RTAFGIEFARSAFQKGRKGAKKVMIVITDGSHSDSPLEKVIQCSERDNTVYAVAVL 300
Db 241 RTAFGIEFARSAFQKGRKGAKKVMIVITDGSHSDSPLEKVIQCSERDNTVYAVAVL 300

QY 301 GYNRRGINPETELNEIKYIASDPDDKHFFNFVTDAAALKDIDVALGDRIFSLGTTNKNET 360
Db 301 GYNRRGINPETELNEIKYIASDPDDKHFFNFVTDAAALKDIDVALGDRIFSLGTTNKNET 360
QY 361 SFGLEMSQTFSSHVVEDGVLLGAVGAYDWNGAVLAKETSAGKVIPLRESYLKEFPPELKN 420
Db 361 SFGLEMSQTFSSHVVEDGVLLGAVGAYDWNGAVLAKETSAGKVIPLRESYLKEFPPELKN 420
QY 421 HGAYLGYTIVTSVSSRQGRVYVAGAPRFNHTGKVILFTMHNNRSLTIHQMRRGQQTGSYF 480
Db 421 HGAYLGYTIVTSVSSRQGRVYVAGAPRFNHTGKVILFTMHNNRSLTIHQMRRGQQTGSYF 480
QY 481 GSIITSVDIDGCVTDVLLVGAAPMYNEGERGKVVYELRQNRVFNGLTKDSHSYQNA 540
Db 481 GSIITSVDIDGCVTDVLLVGAAPMYNEGERGKVVYELRQNRVFNGLTKDSHSYQNA 540
QY 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPKORITASELATG 600
Db 541 RFGSSIASVRDLNODSYNDVVVGAPLEDNHAGAIYIFHGFGRGSIILKTPKORITASELATG 600
QY 601 LQYFGCSIHGQDLNEDGLIDLAVALGNAVILWSRPVQINASLHFEPSKINIFHRDCK 660
Db 601 LQYFGCSIHGQDLNEDGLIDLAVALGNAVILWSRPVQINASLHFEPSKINIFHRDCK 660
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIYNATMDERRVTPRAHLDEGDRFTNRAV 720
Db 661 RSGRDATCLAAFLCFTPIFLAPHFQTTVGIYNATMDERRVTPRAHLDEGDRFTNRAV 720
QY 721 LLSGGQELCERINFHVLDADYVVPVTSFVSVEYSLSDPDHGPMLDDGWPPTTLRVSVYPPFWG 780
Db 721 LLSGGQELCERINFHVLDADYVVPVTSFVSVEYSLSDPDHGPMLDDGWPPTTLRVSVYPPFWG 780
QY 781 CNEDEHCVPDLVLDARSDDLPTAMEYCORVLRKPAQDCSAYTILSPDTTTFIESTRQAV 840
Db 781 CNEDEHCVPDLVLDARSDDLPTAMEYCORVLRKPAQDCSAYTILSPDTTTFIESTRQAV 840
QY 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERRLQKQVCNVSYPFF 900
Db 841 EATLENRGENAYSTVLNISQSANLQFASLIQKEDSDGSECVNEERRLQKQVCNVSYPFF 900
QY 901 RAKAKVAFRLD 911
Db 901 RAKAKVAFRLD 911

RESULT 5

US-10-262-839-4
; Sequence 4, Application US/10262839
; Publication No. US20040038877A1
; GENERAL INFORMATION:

; APPLICANT: Alsebrook, John,
; APPLICANT: Anderson, David W.,
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Burgess, Catherine,
; APPLICANT: Catterton, Elina,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Ellerman, Karen,
; APPLICANT: Gerlach, Valerie,
; APPLICANT: Gorman, Linda,
; APPLICANT: Guo, Xiaojia,
; APPLICANT: Ji, Weizhen,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Leach, Martin,
; APPLICANT: Li, Li,
; APPLICANT: Miller, Charles,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Reiger, Daniel,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Smithson, Glenda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Taupier, Raymond, jr.,

Db 901 RAKAKVAFRLD 911

|||||

RESULT 8

US-10-262-839-6

; Sequence 6, Application US/10262839

; Publication No. US20040038977A1

; GENERAL INFORMATION:

; APPLICANT: Alsobrook, John,

; APPLICANT: Anderson, David W.,

; APPLICANT: Boldog, Ferenc,

; APPLICANT: Burgess, Catherine,

; APPLICANT: Catterton, Elina,

; APPLICANT: Edinger, Shlomit,

; APPLICANT: Ellerman, Karen,

; APPLICANT: Gerlach, Valerie,

; APPLICANT: Gorman, Linda,

; APPLICANT: Guo, Xiaojia,

; APPLICANT: Ji, Weizhen,

; APPLICANT: Kekuda, Ramesh,

; APPLICANT: Leach, Martin,

; APPLICANT: Li, Li,

; APPLICANT: Miller, Charles,

; APPLICANT: Patturajan, Meera,

; APPLICANT: Reiger, Daniel,

; APPLICANT: Rotenberg, Mark,

; APPLICANT: Shimkets, Richard,

; APPLICANT: Smithson, Glennda,

; APPLICANT: Spytek, Kimberly,

; APPLICANT: Taupier, Raymond, jr.,

; APPLICANT: Vernet, Corine,

; APPLICANT: Voss, Edward,

; APPLICANT: Zerhusen, Brian,

; APPLICANT: Zhong, Mei

; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHOD

; FILE REFERENCE: 21402-462A

; CURRENT APPLICATION NUMBER: US/10/262,839

; CURRENT FILING DATE: 2002-10-01

; PRIOR APPLICATION NUMBER: 60/326,483

; PRIOR FILING DATE: 2001-10-02

; PRIOR APPLICATION NUMBER: 60/327,917

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/328,029

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/328,056

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/381,101

; PRIOR FILING DATE: 2002-05-16

; PRIOR APPLICATION NUMBER: 60/371,972

; PRIOR FILING DATE: 2002-04-12

; PRIOR APPLICATION NUMBER: 60/327,342

; PRIOR FILING DATE: 2001-10-05

; PRIOR APPLICATION NUMBER: 60/328,044

; PRIOR FILING DATE: 2001-10-09

; PRIOR APPLICATION NUMBER: 60/328,849

; PRIOR FILING DATE: 2001-10-12

; PRIOR APPLICATION NUMBER: 60/374,738

; PRIOR FILING DATE: 2002-04-23

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 367

; SOFTWARE: Curaseq1st version 0.1

; SEQ ID NO 6

; LENGTH: 1120

; TYPE: PRT

; ORGANISM: Homo sapiens

US-10-262-839-6

Query Match 63.4%; Score 753; DB 12; Length 1120;

Best Local Similarity 100.0%; Pred. No. 0;

Matches 753; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 159 CQYMDIVIVLDSGNSIYPWVEVQHFLINILKKFYLGPGQIQGVVQYGEDVVFHFLND 218

Db 90 CQYMDIVIVLDSGNSIYPWVEVQHFLINILKKFYLGPGQIQGVVQYGEDVVFHFLND 149

QY 219 YRSVKDWEAAASHIEQRCGTETRTAFGIEFARSEAFQKGRGAKAKVMIVITDGEHSDP 278

Db 150 YRSVKDWEAAASHIEQRCGTETRTAFGIEFARSEAFQKGRGAKAKVMIVITDGEHSDP 209

QY 279 DLEKVIQOESRDNVTRYAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTEAAL 338

Db 210 DLEKVIQOESRDNVTRYAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTEAAL 269

QY 339 KDIVDALGDRIFSLGNTNKNETSFGLEMSQTFSSHVVEDGVLLGAVGAYDMNGAVLKET 398

Db 270 KDIVDALGDRIFSLGNTNKNETSFGLEMSQTFSSHVVEDGVLLGAVGAYDMNGAVLKET 329

QY 399 SAGKVIPLRESYLKEPPEELKNHGAIGLVTVTSVSSROGRVYVACAPFNHTGKILFT 458

Db 330 SAGKVIPLRESYLKEPPEELKNHGAIGLVTVTSVSSROGRVYVACAPFNHTGKILFT 389

QY 459 MHNRSLTTHQAMRGQOIGSYFGSEITSVDIDGQVTDVLLVGAPMYFNEGREGKVVYV 518

Db 390 MHNRSLTTHQAMRGQOIGSYFGSEITSVDIDGQVTDVLLVGAPMYFNEGREGKVVYV 449

QY 519 ELQRNPVYNGTLKDSHSYQNARFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFH 578

Db 450 ELQRNPVYNGTLKDSHSYQNARFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFH 509

QY 579 GFSGSILKTPKQRTASELATGLQYFGCSIHGOLDNEDGLIDLAVGALGNAILWSRPV 638

Db 510 GFSGSILKTPKQRTASELATGLQYFGCSIHGOLDNEDGLIDLAVGALGNAILWSRPV 569

QY 639 VQINASLHFEPSKINIFHRDCKSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMD 698

Db 570 VQINASLHFEPSKINIFHRDCKSGRDATCLAAFLCFTPIFLAPHFQTTTGVIRYNATMD 629

QY 699 ERRYTPRAHLDEGGDRFTNRAVILLSGQELCERINHFHVLDTADYVKEPTVESYSLDDPD 758

Db 630 ERRYTPRAHLDEGGDRFTNRAVILLSGQELCERINHFHVLDTADYVKEPTVESYSLDDPD 689

QY 759 HGPMLDDGWPTTLRVSVVPFWNGCNEDEHCVDPDLVDARSDLPTAMEYQCVLRKPAQDCS 818

Db 690 HGPMLDDGWPTTLRVSVVPFWNGCNEDEHCVDPDLVDARSDLPTAMEYQCVLRKPAQDCS 749

QY 819 AYTLSPDTTTFIESTRQVAVATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGS 878

Db 750 AYTLSPDTTTFIESTRQVAVATLENRGENAYSTVLNISQSANLOFASLIQKEDSDGS 809

QY 879 IECVNEERLQKQVCNVSVYPPFRKAKVAFRLD 911

Db 810 IECVNEERLQKQVCNVSVYPPFRKAKVAFRLD 842

RESULT 9

US-09-764-870-313

; Sequence 313, Application US/09764870

; Patent No. US20020042386A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PTZ14

; CURRENT APPLICATION NUMBER: US/09/764,870

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - consult PALM or file wrapper

; NUMBER OF SEQ ID NOS: 646

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 313

; LENGTH: 707

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-764-870-313

Query Match 41.2%; Score 489; DB 9; Length 707;

Best Local Similarity 99.7%; Pred. No. 0;

Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MDLPRLGVAVAWALSMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOHDISGNKMLVGA 60
Db 17 MDLPRLGVAVAWALSMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOHDISGNKMLVGA 76
QY 61 PLETNQYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKDNMRGLSLATNPKNSFLA 120
Db 77 PLETNQYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKDNMRGLSLATNPKNSFLA 136
QY 121 CSPLWSEHCSSYYTTGMSRVNSNFRFSTKVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db 137 CSPLWSEHCSSYYTTGMSRVNSNFRFSTKVAPALQRCQTYMDIVIVLDGNSIYPWVE 196
QY 181 VQFLNINILKFFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 240
Db 197 VQFLNINILKFFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 256
QY 241 RTAFGIEFARSEAFQKGRKGAKVMIVITDGESHSDPLEKVIQOESRDNVTRYAVAVL 300
Db 257 RTAFGIEFARSEAFQKGRKGAKVMIVITDGESHSDPLEKVIQOESRDNVTRYAVAVL 316
QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLEGTNKNET 360
Db 317 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLEGTNKNET 376
QY 361 SFGLEMSQTFSSHVVVEDGVLGAVGAYDNGAVLKETSAKVIPLRESYLKEFPEELKN 420
Db 377 SFGLEMSQTFSSHVVVEDGVLGAVGAYDNGAVLKETSAKVIPLRESYLKEFPEELKN 436
QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKVLFTVHNNRSLTIHQMRGQQIGSYF 480
Db 437 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKVLFTVHNNRSLTIHQMRGQQIGSYF 496
QY 481 GSEITSDVDGCGVTDVLLVGPMPYFNEGRERGVVYVELRQNRVYNGTLKDSHSYQNA 540
Db 497 GSEITSDVDGCGVTDVLLVGPMPYFNEGRERGVVYVELRQNRVYNGTLKDSHSYQNA 556
QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 600
Db 557 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 616
QY 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQVQINASHLHFEPSKINIFHRDCK 660
Db 617 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQVQINASHLHFEPSKINIFHRDCK 676
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 691
Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 707
```

RESULT 10

US-10-125-540-313
; Sequence 313, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ:4C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; CURRENT FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 313
; LENGTH: 707
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-125-540-313

Query Match 41.2%; Score 489; DB 14; Length 707;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 689; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 MDLPRLGVAVAWALSMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOHDISGNKMLVGA 60
Db 17 MDLPRLGVAVAWALSMPGFTDTFNMTRKPRVTPGSRRTAFPGYTVQOHDISGNKMLVGA 76
QY 61 PLETNQYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKDNMRGLSLATNPKNSFLA 120
Db 77 PLETNQYQKTGDYVKCPVHGNCTKLNLRVTLSNVSEKDNMRGLSLATNPKNSFLA 136
QY 121 CSPLWSEHCSSYYTTGMSRVNSNFRFSTKVAPALQRCQTYMDIVIVLDGNSIYPWVE 180
Db 137 CSPLWSEHCSSYYTTGMSRVNSNFRFSTKVAPALQRCQTYMDIVIVLDGNSIYPWVE 196
QY 181 VQFLNINILKFFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 240
Db 197 VQFLNINILKFFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQGGTET 256
QY 241 RTAFGIEFARSEAFQKGRKGAKVMIVITDGESHSDPLEKVIQOESRDNVTRYAVAVL 300
Db 257 RTAFGIEFARSEAFQKGRKGAKVMIVITDGESHSDPLEKVIQOESRDNVTRYAVAVL 316
QY 301 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLEGTNKNET 360
Db 317 GYNNRGINPETFLNEIKYIASDPDDKHFFNVTDAAALKDVIDALGDRIFSLEGTNKNET 376
QY 361 SFGLEMSQTFSSHVVVEDGVLGAVGAYDNGAVLKETSAKVIPLRESYLKEFPEELKN 420
Db 377 SFGLEMSQTFSSHVVVEDGVLGAVGAYDNGAVLKETSAKVIPLRESYLKEFPEELKN 436
QY 421 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKVLFTVHNNRSLTIHQMRGQQIGSYF 480
Db 437 HGAYLGYTVTSVSSRQGRVYVAGAPRNFHTKVLFTVHNNRSLTIHQMRGQQIGSYF 496
QY 481 GSEITSDVDGCGVTDVLLVGPMPYFNEGRERGVVYVELRQNRVYNGTLKDSHSYQNA 540
Db 497 GSEITSDVDGCGVTDVLLVGPMPYFNEGRERGVVYVELRQNRVYNGTLKDSHSYQNA 556
QY 541 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 600
Db 557 RFGSSIASVRDLNODSYNDVVGAPLEDNHAGAIYIFHGRGSIILKTPKQITASELATG 616
QY 601 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQVQINASHLHFEPSKINIFHRDCK 660
Db 617 LQYFGCSIHQDLNEDGLIDLAVGALGNVILWSPVQVQINASHLHFEPSKINIFHRDCK 676
QY 661 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 691
Db 677 RSGRDATCLAAFLCFTPIFLAPHFQTTTGI 707
```

RESULT 11

US-10-108-260A-3386
; Sequence 3386, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
; APPLICANT: HELIX RESEARCH INSTITUTE
; TITLE OF INVENTION: No US20040005560A1el full length cDNA
; FILE REFERENCE: H1-A0106
; CURRENT APPLICATION NUMBER: US/10/108,260A
; CURRENT FILING DATE: 2002-03-27
; NUMBER OF SEQ ID NOS: 5458
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3386
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-108-260A-3386

Query Match 20.6%; Score 245; DB 15; Length 437;
Best Local Similarity 100.0%; Pred. No. 1e-224;
Matches 245; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 175 IYPVWEVQHFNLINILKFFYIGPQIQVGVQYGEDVVEHFLNDYRSVKDVVEAASHIEQ 234

Db 175 IYPWVEVQHFLNLIKFFYIGPGQIQGVVQGVGVHFEHLNDYRSVKDVVEAASHIEQ 234
QY 235 RGGTETRTAFGIEFARSEAFQGRGKAKKVMIVITDGESHDPDLEKVIQOSERDNTVR 294
Db 235 RGGTETRTAFGIEFARSEAFQGRGKAKKVMIVITDGESHDPDLEKVIQOSERDNTVR 294
QY 295 YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIFSLG 354
Db 295 YAVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAALKDIDVALGDRIFSLG 354
QY 355 TNKNETSFGLNSQTFSSHVEDGYLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEF 414
Db 355 TNKNETSFGLNSQTFSSHVEDGYLLGAVGAYDNGAVLKETSAGKVIPLRESYLKEF 414
QY 415 PEEK 419
Db 415 PEEK 419

RESULT 12
US-09-805-354-8
; Sequence 8, Application US/09805354
; Publication No. US20030078375A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-536001 US/09/805,354
; CURRENT APPLICATION NUMBER: US/09/805,354
; PRIOR FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-354-8

Query Match 16.2%; Score 193; DB 10; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.2e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 159 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGPGQIQGVVQGVGVHFEHLND 218
Db 1 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGPGQIQGVVQGVGVHFEHLND 60
QY 219 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQGRGKAKKVMIVITDGESHDP 278
Db 61 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQGRGKAKKVMIVITDGESHDP 120
QY 279 DLEKVIQOSERDNTVRVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 338
Db 121 DLEKVIQOSERDNTVRVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 180
QY 339 KDIVDALGDRIFS 351
Db 181 KDIVDALGDRIFS 193

RESULT 13
US-09-758-493-8
; Sequence 8, Application US/09758493
; Publication No. US20040086935A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping

; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 00786-804001
; CURRENT APPLICATION NUMBER: US/09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-493-8

Query Match 16.2%; Score 193; DB 12; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.2e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 159 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGPGQIQGVVQGVGVHFEHLND 218
Db 1 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGPGQIQGVVQGVGVHFEHLND 60
QY 219 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQGRGKAKKVMIVITDGESHDP 278
Db 61 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQGRGKAKKVMIVITDGESHDP 120
QY 279 DLEKVIQOSERDNTVRVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 338
Db 121 DLEKVIQOSERDNTVRVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 180
QY 339 KDIVDALGDRIFS 351
Db 181 KDIVDALGDRIFS 193

RESULT 14
US-10-144-259-8
; Sequence 8, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 193
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-8

Query Match 16.2%; Score 193; DB 14; Length 193;
Best Local Similarity 100.0%; Pred. No. 2.2e-175;
Matches 193; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 159 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGPGQIQGVVQGVGVHFEHLND 218
Db 1 CQYMDIVIVLDGNSIYPWVEVQHFLNLIKFFYIGPGQIQGVVQGVGVHFEHLND 60
QY 219 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQGRGKAKKVMIVITDGESHDP 278
Db 61 YRSVKDVVEAASHIEQGGTETRTAFGIEFARSEAFQGRGKAKKVMIVITDGESHDP 120
QY 279 DLEKVIQOSERDNTVRVAVLGYNNRGINPETFLNEIKYIASDPDDKHFFNVTDEAAL 338

Db 121 LDKVIOQSDNTRYAVAVLGYNRRGINPFTFLNEIKYIASDDDKHFFNVDEAAL 180
|||||

Qy 339 KDIVDALGDRIFS 351
|||||
Db 181 KDIVDALGDRIFS 193
|||||

RESULT 15

US-09-764-870-472
; Sequence 472, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 472
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-870-472

Query Match 7.4%; Score 88; DB 9; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.4e-75;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLND 218
|||||

Db 8 CQYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLND 67
|||||

Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246
|||||
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95
|||||

RESULT 16

US-09-764-875-1193
; Sequence 1193, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PUZ02
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1193
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-875-1193

Query Match 7.4%; Score 88; DB 11; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.4e-75;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLND 218
|||||

Db 8 CQYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLND 67
|||||

Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246
|||||
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95
|||||

RESULT 17

US-10-125-540-472
; Sequence 472, Application US/10125540
; Publication No. US20030059875A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14C1
; CURRENT APPLICATION NUMBER: US/10/125,540
; CURRENT FILING DATE: 2002-04-19
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 472
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-125-540-472

Query Match 7.4%; Score 88; DB 14; Length 103;
Best Local Similarity 100.0%; Pred. No. 2.4e-75;
Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 159 CQYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLND 218
|||||

Db 8 CQYMDIVIVLDGNSIYPWVEVQHFLINILKKFYIGPGQIQGVVQYGEDVVHFEHLND 67
|||||

Qy 219 YRSVKDVEAAASHIEQGGTETRTAFGI 246
|||||
Db 68 YRSVKDVEAAASHIEQGGTETRTAFGI 95
|||||

RESULT 18

US-10-144-259-25
; Sequence 25, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 42
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-25

Query Match 3.5%; Score 42; DB 14; Length 42;
Best Local Similarity 100.0%; Pred. No. 9.2e-32;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 335 EAAKDIVDALGDRIPSLSGTNKNETSPGLEMSQTGFSSHV 376
|||||

```
Db 1 EAAKDIVDALGDRIFSLGKTNKNETSPGLEMSQTGFSSHVV 42

RESULT 19
US-09-866-050A-500
; Sequence 500, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 500
; LENGTH: 545
; TYPE: PRT
; ORGANISM: Mouse
US-09-866-050A-500

Query Match 3.5%; Score 42; DB 10; Length 545;
Best Local Similarity 100.0%; Pred. No. 9.1e-31;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 807
|||||
Db 123 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 164
|||||

RESULT 20
US-09-866-050A-624
; Sequence 624, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 624
; LENGTH: 688
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-624

Query Match 3.5%; Score 42; DB 10; Length 688;
Best Local Similarity 100.0%; Pred. No. 1.1e-30;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 807
|||||
Db 266 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 307
|||||

RESULT 21
US-09-866-050A-501
; Sequence 501, Application US/09866050A
; Publication No. US20030040471A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; TITLE OF INVENTION: and Methods for Their Use
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 501
; LENGTH: 696
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-501

Query Match 3.5%; Score 42; DB 10; Length 696;
Best Local Similarity 100.0%; Pred. No. 1.1e-30;
Matches 42; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 766 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 807
|||||
Db 266 GWPTTLRVSPFWNGCNEDEHCVDPDLVLDARSDLPTAMEYQC 307
|||||

RESULT 22
US-09-984-130-102
; Sequence 102, Application US/09984130
; Publication No. US2003005231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 102
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-102

Query Match 2.8%; Score 33; DB 10; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.9e-23;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 64 TNGYQKTGDIVYKCPVHGNCCTKLNLRVTLSNV 96
|||||
Db 1 TNGYQKTGDIVYKCPVHGNCCTKLNLRVTLSNV 33
|||||

RESULT 23
US-09-836-353A-102
; Sequence 102, Application US/09836353A
; Publication No. US20030129685A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
```

```

; FILE REFERENCE: PF489P1
; CURRENT APPLICATION NUMBER: US/09/836,353A
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 147
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 102
; LENGTH: 33
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-836-353A-102

Query Match      2.8%; Score 33; DB 10; Length 33;
Best Local Similarity 100.0%; Pred. No. 2.9e-23;
Matches 33; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      64 TNGYQXTGDKVYKCPVHGNCTKLNIGRVTLSNV 96
Db      1 TNGYQXTGDKVYKCPVHGNCTKLNIGRVTLSNV 33

RESULT 24
US-09-866-050A-391
; Sequence 391, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 391
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-391

Query Match      1.8%; Score 21; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.4e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      636 RPVVQINASLHPEPSKINIFH 656
Db      57 RPVVQINASLHPEPSKINIFH 77

RESULT 25
US-09-866-050A-390
; Sequence 390, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Strachan, Lorna
; APPLICANT: Sleeman, Matthew
; APPLICANT: Onrust, Rene
; APPLICANT: Murison, James G.
; APPLICANT: Kumble, Krishanand D.
; TITLE OF INVENTION: Compositions Isolated From Skin Cells
; FILE REFERENCE: 11000.1011c4U
; CURRENT APPLICATION NUMBER: US/09/866,050A
; CURRENT FILING DATE: 2001-05-24
; NUMBER OF SEQ ID NOS: 725
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 390
; LENGTH: 158
; TYPE: PRT
; ORGANISM: Rat
US-09-866-050A-390

Query Match      1.8%; Score 21; DB 10; Length 158;
Best Local Similarity 100.0%; Pred. No. 3.4e-11;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      636 RPVVQINASLHPEPSKINIFH 656
Db      57 RPVVQINASLHPEPSKINIFH 77

RESULT 26
US-09-764-870-284
; Sequence 284, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 284
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-284

Query Match      1.3%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVILDGNSIYPW 178
Db      47 DIVILDGNSIYPW 61

RESULT 27
US-09-764-870-444
; Sequence 444, Application US/09764870
; Patent No. US20020042386A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ14
; CURRENT APPLICATION NUMBER: US/09/764,870
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 646
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 444
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-870-444

Query Match      1.3%; Score 15; DB 9; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.7e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVILDGNSIYPW 178
Db      47 DIVILDGNSIYPW 61

```

RESULT 28

US-10-125-540-284
 ; Sequence 284, Application US/10125540
 ; Publication No. US20030059875A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT214C1
 ; CURRENT APPLICATION NUMBER: US/10/125,540
 ; PRIOR FILING DATE: 2002-04-19
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 646
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 284
 ; LENGTH: 148
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-125-540-284

Query Match 1.3%; Score 15; DB 14; Length 148;
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
 |||||
 Db 47 DIVIVLDGNSIYPW 61

RESULT 29

US-10-125-540-444
 ; Sequence 444, Application US/10125540
 ; Publication No. US20030059875A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT214C1
 ; CURRENT APPLICATION NUMBER: US/10/125,540
 ; PRIOR FILING DATE: 2002-04-19
 ; Prior Application removed - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 646
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 444
 ; LENGTH: 148
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-125-540-444

Query Match 1.3%; Score 15; DB 14; Length 148;
 Best Local Similarity 100.0%; Pred. No. 1.7e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
 |||||
 Db 47 DIVIVLDGNSIYPW 61

RESULT 30

US-10-346-863-57
 ; Sequence 57, Application US/10346863
 ; Publication No. US20040038325A1
 ; GENERAL INFORMATION:
 ; APPLICANT: PHELPS, CHRISTOPHER BENJAMIN
 ; APPLICANT: FAGAN, RICHARD JOSEPH
 ; APPLICANT: GUTTERIDGE, ALEX
 ; TITLE OF INVENTION: ADHESION MOLECULES
 ; FILE REFERENCE: 674575-2001
 ; CURRENT APPLICATION NUMBER: US/10/346,863
 ; CURRENT FILING DATE: 2003-01-17
 ; PRIOR APPLICATION NUMBER: PCT/GB01/03318
 ; PRIOR FILING DATE: 2001-07-24
 ; PRIOR APPLICATION NUMBER: GB 0018126.3

; PRIOR FILING DATE: 2000-07-24
 ; PRIOR APPLICATION NUMBER: GB 0025447.4
 ; PRIOR FILING DATE: 2000-10-17
 ; NUMBER OF SEQ ID NOS: 57
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 57
 ; LENGTH: 192
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-10-346-863-57

Query Match 1.3%; Score 15; DB 12; Length 192;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
 |||||
 Db 5 DIVIVLDGNSIYPW 19

RESULT 31

US-10-474-832-59
 ; Sequence 59, Application US/10474832
 ; Publication No. US20040081651A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BIOGEN, INC.
 ; TITLE OF INVENTION: ANTIBODIES TO VLA-1
 ; FILE REFERENCE: A101 PCT
 ; CURRENT APPLICATION NUMBER: US/10/474,832
 ; CURRENT FILING DATE: 2003-10-14
 ; PRIOR APPLICATION NUMBER: 60/283,794
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: 60/303,689
 ; PRIOR FILING DATE: 2001-07-06
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 59
 ; LENGTH: 192
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: Rat/human chimeric
 ; OTHER INFORMATION: I domain construct
 US-10-474-832-59

Query Match 1.3%; Score 15; DB 16; Length 192;
 Best Local Similarity 100.0%; Pred. No. 2.2e-05;
 Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 DIVIVLDGNSIYPW 178
 |||||
 Db 4 DIVIVLDGNSIYPW 18

RESULT 32

US-10-474-832-60
 ; Sequence 60, Application US/10474832
 ; Publication No. US20040081651A1
 ; GENERAL INFORMATION:
 ; APPLICANT: BIOGEN, INC.
 ; TITLE OF INVENTION: ANTIBODIES TO VLA-1
 ; FILE REFERENCE: A101 PCT
 ; CURRENT APPLICATION NUMBER: US/10/474,832
 ; CURRENT FILING DATE: 2003-10-14
 ; PRIOR APPLICATION NUMBER: 60/283,794
 ; PRIOR FILING DATE: 2001-04-13
 ; PRIOR APPLICATION NUMBER: 60/303,689
 ; PRIOR FILING DATE: 2001-07-06
 ; NUMBER OF SEQ ID NOS: 70
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 60
 ; LENGTH: 192
 ; TYPE: PRT

```

; ORGANISM: Rattus sp.
US-10-474-832-60

Query Match      1.3%; Score 15; DB 16; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      4 DIVVLGDSNSIYPW 18

RESULT 33
US-10-474-832-61
; Sequence 61, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 61
; LENGTH: 192
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-474-832-61

Query Match      1.3%; Score 15; DB 16; Length 192;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      4 DIVVLGDSNSIYPW 18

RESULT 34
US-09-805-354-5
; Sequence 5, Application US/09805354
; Publication No. US20030078375A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-536001
; CURRENT APPLICATION NUMBER: US/09/805,354
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-354-5

Query Match      1.3%; Score 15; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      4 DIVVLGDSNSIYPW 178

```

```

Db      6 DIVVLGDSNSIYPW 20

RESULT 35
US-09-758-493-5
; Sequence 5, Application US/09758493
; Publication No. US20040086935A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 00786-804001
; CURRENT APPLICATION NUMBER: US/09/758,493
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-493-5

Query Match      1.3%; Score 15; DB 12; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      6 DIVVLGDSNSIYPW 20

RESULT 36
US-10-144-259-5
; Sequence 5, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-5

Query Match      1.3%; Score 15; DB 14; Length 195;
Best Local Similarity 100.0%; Pred. No. 2.2e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYPW 178
Db      6 DIVVLGDSNSIYPW 20

RESULT 37
US-09-996-738-5
; Sequence 5, Application US/09996738
; Patent No. US20020146417A1
; GENERAL INFORMATION:

```

```
; APPLICANT: Biogen, Inc.
; APPLICANT: De Fougères, Antonin
; APPLICANT: Gotwals, Philip
; APPLICANT: Lobb, Roy
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Inflammatory
; TITLE OF INVENTION: Disorders
; FILE REFERENCE: A076PCT
; CURRENT APPLICATION NUMBER: US/09/996,738
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/185336
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/137038
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 212
; TYPE: PRT
; ORGANISM: rat
US-09-996-738-5
```

```
Query Match 1.3%; Score 15; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36
```

RESULT 38

```
US-09-996-738-6
; Sequence 6, Application US/09996738
; Patent No. US20020146417A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: De Fougères, Antonin
; APPLICANT: Gotwals, Philip
; APPLICANT: Lobb, Roy
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Inflammatory
; TITLE OF INVENTION: Disorders
; FILE REFERENCE: A076PCT
; CURRENT APPLICATION NUMBER: US/09/996,738
; CURRENT FILING DATE: 2001-11-30
; PRIOR APPLICATION NUMBER: 60/185336
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: 60/137038
; PRIOR FILING DATE: 1999-06-01
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 212
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-996-738-6
```

```
Query Match 1.3%; Score 15; DB 9; Length 212;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36
```

RESULT 39

```
US-10-625-260-5
; Sequence 5, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
```

```
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rat
US-10-625-260-5
```

```
Query Match 1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36
```

RESULT 40

```
US-10-625-260-6
; Sequence 6, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-625-260-6
```

```
Query Match 1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 164 DIVVLGSGNSIYPW 178
Db 22 DIVVLGSGNSIYPW 36
```

RESULT 41

```
US-10-625-260-9
; Sequence 9, Application US/10625260
; Publication No. US20040037827A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliansky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/625,260
; CURRENT FILING DATE: 2003-07-22
; PRIOR APPLICATION NUMBER: 60/137,214
```

```
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-625-260-9

Query Match      1.3%; Score 15; DB 12; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178
Db      22 DIVILDGNSIYPW 36

RESULT 42
US-10-061-658-5
; Sequence 5, Application US/10061658
; Publication No. US20020182213A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rat
US-10-061-658-5

Query Match      1.3%; Score 15; DB 13; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178
Db      22 DIVILDGNSIYPW 36

RESULT 43
US-10-061-658-6
; Sequence 6, Application US/10061658
; Publication No. US20020182213A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rat
US-10-061-658-6

Query Match      1.3%; Score 15; DB 13; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178
Db      22 DIVILDGNSIYPW 36

RESULT 44
US-10-061-658-9
; Sequence 9, Application US/10061658
; Publication No. US20020182213A1
; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.
; APPLICANT: Gotwals, Philip
; APPLICANT: Koteliarsky, Victor
; TITLE OF INVENTION: Method for the Treatment of Fibrosis
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-061-658-9

Query Match      1.3%; Score 15; DB 13; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178
Db      22 DIVILDGNSIYPW 36

RESULT 45
US-10-474-832-63
; Sequence 63, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCI
; CURRENT APPLICATION NUMBER: US/10/474,832
; CURRENT FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 63
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Rattus sp.
US-10-474-832-63

Query Match      1.3%; Score 15; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      164 DIVILDGNSIYPW 178
```



```
Db      22 DIVVLGSGNSIYFW 36
|||||
RESULT 46
US-10-474-832-64
; Sequence 64, Application US/10474832
; Publication No. US20040081651A1
; GENERAL INFORMATION:
; APPLICANT: BIOGEN, INC.
; TITLE OF INVENTION: ANTIBODIES TO VLA-1
; FILE REFERENCE: A101 PCT
; CURRENT APPLICATION NUMBER: US/10/474,832
; PRIOR FILING DATE: 2003-10-14
; PRIOR APPLICATION NUMBER: 60/283,794
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/303,689
; PRIOR FILING DATE: 2001-07-06
; NUMBER OF SEQ ID NOS: 70
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-474-832-64

Query Match      1.3%; Score 15; DB 16; Length 214;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYFW 178
|||||
Db      22 DIVVLGSGNSIYFW 36
|||||

RESULT 47
US-09-984-130-103
; Sequence 103, Application US/09984130
; Publication No. US2003005231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-103

Query Match      1.3%; Score 15; DB 10; Length 1151;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYFW 178
|||||
Db      144 DIVVLGSGNSIYFW 158
|||||

RESULT 48
US-09-984-130-103
; Sequence 103, Application US/09984130
; Publication No. US2003005231A1
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 12 Human Secreted Proteins
; FILE REFERENCE: PF489P2
; CURRENT APPLICATION NUMBER: US/09/984,130
; CURRENT FILING DATE: 2001-10-29
; PRIOR APPLICATION NUMBER: 60/243,792
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: 09/836,353
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/198,407
; PRIOR FILING DATE: 2000-04-19
; PRIOR APPLICATION NUMBER: PCT/US99/25031
; PRIOR FILING DATE: 1999-10-27
; PRIOR APPLICATION NUMBER: 60/105,971
; PRIOR FILING DATE: 1998-10-28
; NUMBER OF SEQ ID NOS: 149
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 103
; LENGTH: 1151
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-984-130-103

Query Match      1.3%; Score 15; DB 10; Length 1151;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYFW 178
|||||
Db      144 DIVVLGSGNSIYFW 158
|||||

RESULT 49
US-09-918-715-250
; Sequence 250, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
; APPLICANT: Best Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107.00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 250
; LENGTH: 1179
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-918-715-250

Query Match      1.3%; Score 15; DB 12; Length 1179;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGSGNSIYFW 178
|||||
Db      172 DIVVLGSGNSIYFW 186
|||||

RESULT 50
US-09-918-715-307
; Sequence 307, Application US/09918715
; Publication No. US20030017157A1
; GENERAL INFORMATION:
; APPLICANT: Brad St. Croix
```

```
; APPLICANT: Bert Vogelstein
; APPLICANT: Kenneth Kinzler
; TITLE OF INVENTION: ENDOTHELIAL CELL EXPRESSION PATTERNS
; FILE REFERENCE: 1107,00134
; CURRENT APPLICATION NUMBER: US/09/918,715
; CURRENT FILING DATE: 2001-08-01
; PRIOR APPLICATION NUMBER: 60/222,599
; PRIOR FILING DATE: 2000-08-02
; PRIOR APPLICATION NUMBER: 60/224,360
; PRIOR FILING DATE: 2000-08-11
; PRIOR APPLICATION NUMBER: 60/282,850
; PRIOR FILING DATE: 2000-04-11
; NUMBER OF SEQ ID NOS: 358
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 307
; LENGTH: 1180
; TYPE: PRT
; ORGANISM: Rat
US-09-918-715-307
```

```
Query Match 1.1%; Score 15; DB 12; Length 1180;
Best Local Similarity 100.0%; Pred. No. 0.00011;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 164 DIVVLGDSNSIYPW 178
Db 172 DIVVLGDSNSIYPW 186
|||||
```

```
RESULT 51
US-09-805-354-7
; Sequence 7, Application US/09805354
; Publication No. US20030078375A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-536001
; CURRENT APPLICATION NUMBER: US/09/805,354
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-805-354-7
```

```
Query Match 1.1%; Score 13; DB 10; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 166 DIVVLGDSNSIYPW 178
Db 8 DIVVLGDSNSIYPW 20
|||||
```

```
RESULT 52
US-09-758-493-7
; Sequence 7, Application US/09758493
; Publication No. US20040086935A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: HIGH AFFINITY INTEGRIN POLYPEPTIDES AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: 00786-804001
```

```
; CURRENT APPLICATION NUMBER: US/09/758,493
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-758-493-7
```

```
Query Match 1.1%; Score 13; DB 12; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 166 DIVVLGDSNSIYPW 178
Db 8 DIVVLGDSNSIYPW 20
|||||
```

```
RESULT 53
US-10-144-259-7
; Sequence 7, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 195
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-7
```

```
Query Match 1.1%; Score 13; DB 14; Length 195;
Best Local Similarity 100.0%; Pred. No. 0.0018;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 166 DIVVLGDSNSIYPW 178
Db 8 DIVVLGDSNSIYPW 20
|||||
```

```
RESULT 54
US-10-741-601-531
; Sequence 531, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0601500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 531
; LENGTH: 1167
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-531
```

Query Match 1.1%; Score 13; DB 16; Length 1167;
Best Local Similarity 100.0%; Pred. No. 0.0089;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSIYPW 178
Db 169 VIVLDGNSIYPW 181

RESULT 55
US-10-741-601-532
; Sequence 532, Application US/10741601
; Publication No. US20040166519A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: CL001500
; CURRENT APPLICATION NUMBER: US/10/741,601
; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 25415
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 532
; LENGTH: 1177
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-741-601-532

Query Match 1.1%; Score 13; DB 16; Length 1177;
Best Local Similarity 100.0%; Pred. No. 0.0089;
Matches 13; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 166 VIVLDGNSIYPW 178
Db 169 VIVLDGNSIYPW 181

RESULT 56
US-10-144-259-26
; Sequence 26, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 09/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-26

Query Match 0.9%; Score 11; DB 14; Length 43;
Best Local Similarity 100.0%; Pred. No. 0.038;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 340 DIVDALGDRIF 350
Db 6 DIVDALGDRIF 16

RESULT 57
US-10-336-603A-42
; Sequence 42, Application US/10336603A

; Publication No. US20040072997A1
; GENERAL INFORMATION:
; APPLICANT: Alsbrook et al.
; TITLE OF INVENTION: THERAPEUTIC POLYPEPTIDES, NUCLEIC ACIDS ENCODING SAME, AND METHO
; FILE REFERENCE: 21402-533A
; CURRENT APPLICATION NUMBER: US/10/336,603A
; CURRENT FILING DATE: 2003-01-03
; PRIOR APPLICATION NUMBER: 09/746,491
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 16/055,569
; PRIOR FILING DATE: 2001-10-26
; NUMBER OF SEQ ID NOS: 169
; SOFTWARE: Curaseq1ist version 0.1
; SEQ ID NO 42
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-336-603A-42

Query Match 0.9%; Score 11; DB 12; Length 1147;
Best Local Similarity 100.0%; Pred. No. 0.71;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
Db 506 TDVLLVGAPMY 516

RESULT 58
US-10-211-462-187
; Sequence 187, Application US/10211462
; Publication No. US20040033495A1
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Angiogenesis, Compositions and
; FILE REFERENCE: 018501-006200US
; CURRENT APPLICATION NUMBER: US/10/211,462
; CURRENT FILING DATE: 2003-02-13
; PRIOR APPLICATION NUMBER: US 09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/791,390
; PRIOR FILING DATE: 2001-02-22
; PRIOR APPLICATION NUMBER: US 60/310,025
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/334,244
; PRIOR FILING DATE: 2001-11-29
; NUMBER OF SEQ ID NOS: 230
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 187
; LENGTH: 1181
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-211-462-187

Query Match 0.9%; Score 11; DB 12; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
Db 506 TDVLLVGAPMY 516

RESULT 59
US-10-160-354-2
; Sequence 2, Application US/10160354
; Publication No. US20030013107A1
; GENERAL INFORMATION:

APPLICANT: Chu, Peter
APPLICANT: Li, Congfen
APPLICANT: Liao, X. Charlene
APPLICANT: Pardo, Jorge
APPLICANT: Rigel Pharmaceuticals, Incorporated
TITLE OF INVENTION: Alpha 2 Integrin: Modulators of Lymphocyte Activation
FILE REFERENCE: 021044-001110US
CURRENT APPLICATION NUMBER: US/10/160,354
CURRENT FILING DATE: 2002-05-30
PRIOR APPLICATION NUMBER: US 60/296,819
PRIOR FILING DATE: 2001-06-07
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 1181
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: wild type alpha 2 integrin, CD49b
US-10-160-354-2

Query Match 0.9%; Score 11; DB 14; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
Db 506 TDVLLVGAPMY 516
|||||

RESULT 60
US-10-295-027-1286
Sequence 1286, Application US/10295027
Publication No. US20030232350A1
GENERAL INFORMATION:
APPLICANT: Afar, Daniel
APPLICANT: Aziz, Natasha
APPLICANT: Ginsberg, Wendy M.
APPLICANT: Gish, Kurt C.
APPLICANT: Glynn, Richard
APPLICANT: Hevezi, Peter A.
APPLICANT: Mack, David H.
APPLICANT: Murray, Richard
APPLICANT: Watson, Susan R.
APPLICANT: Eos Biotechnology, Inc.
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
Methods of Screening for Modulators of Cancer
FILE REFERENCE: 018501-012500US
CURRENT APPLICATION NUMBER: US/10/295,027
CURRENT FILING DATE: 2002-11-13
PRIOR APPLICATION NUMBER: US 09/663,733
PRIOR FILING DATE: 2000-09-15
PRIOR APPLICATION NUMBER: US 60/350,666
PRIOR FILING DATE: 2001-11-13
PRIOR APPLICATION NUMBER: US 60/335,394
PRIOR FILING DATE: 2001-11-15
PRIOR APPLICATION NUMBER: US 60/332,464
PRIOR FILING DATE: 2001-11-21
PRIOR APPLICATION NUMBER: US 60/334,393
PRIOR FILING DATE: 2001-11-29
PRIOR APPLICATION NUMBER: US 60/340,376
PRIOR FILING DATE: 2001-12-14
PRIOR APPLICATION NUMBER: US 60/347,211
PRIOR FILING DATE: 2002-01-08
PRIOR APPLICATION NUMBER: US 60/347,349
PRIOR FILING DATE: 2002-01-10
PRIOR APPLICATION NUMBER: US 60/355,250
PRIOR FILING DATE: 2002-02-08
PRIOR APPLICATION NUMBER: US 60/356,714
PRIOR FILING DATE: 2002-02-13
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 1386
SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 1286
LENGTH: 1181
TYPE: PRT
ORGANISM: Homo sapiens
US-10-295-027-1286

Query Match 0.9%; Score 11; DB 15; Length 1181;
Best Local Similarity 100.0%; Pred. No. 0.73;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 495 TDVLLVGAPMY 505
Db 506 TDVLLVGAPMY 516
|||||

RESULT 61
US-10-108-260A-3415
Sequence 3415, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20040005560A1 full length cDNA
FILE REFERENCE: HL-A0106
CURRENT APPLICATION NUMBER: US/10/108,260A
CURRENT FILING DATE: 2002-03-27
NUMBER OF SEQ ID NOS: 5458
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3415
LENGTH: 164
TYPE: PRT
ORGANISM: Homo sapiens
US-10-108-260A-3415

Query Match 0.8%; Score 9; DB 15; Length 164;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLLVL 1161
Db 99 LLLALLLVL 107
|||||

RESULT 62
US-10-036-854-102
Sequence 102, Application US/10038854
Publication No. US20040022781A1
GENERAL INFORMATION:
APPLICANT: Spytex, Kimberly A
APPLICANT: Li, Li
APPLICANT: Wolenc, Adam R
APPLICANT: Vernet, Corine
APPLICANT: Eisen, Andrew J
APPLICANT: Liu, Xiaohong
APPLICANT: Malyankar, Uriel M
APPLICANT: Shinkets, Richard A
APPLICANT: Tchernev, Velizar
APPLICANT: Spaderna, Steven K
APPLICANT: Gorman, Linda
APPLICANT: Kekuda, Ramesh
APPLICANT: Patturajan, Meera
APPLICANT: Gusev, Vladimir Y
APPLICANT: Gangolli, Bsha A
APPLICANT: Guo, Xiaojia S
APPLICANT: Shenoy, Suresh G
APPLICANT: Rastelli, Luca
APPLICANT: Casman, Stacie J
APPLICANT: Boldog, Ferenc
APPLICANT: Burgess, Catherine E
APPLICANT: Edinger, Shlomit R
APPLICANT: Ellerman, Karen
APPLICANT: Gunther, Erik
APPLICANT: Smithson, Glenda
APPLICANT: Millet, Isabelle

```
; APPLICANT: MacDougall, John R
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-230
; CURRENT APPLICATION NUMBER: US/10/038,854
; CURRENT FILING DATE: 2003-01-22
; PRIOR APPLICATION NUMBER: 60/258,928
; PRIOR FILING DATE: 2000-12-29
; PRIOR APPLICATION NUMBER: 60/259,415
; PRIOR FILING DATE: 2001-01-02
; PRIOR APPLICATION NUMBER: 60/259,785
; PRIOR FILING DATE: 2001-01-04
; PRIOR APPLICATION NUMBER: 60/269,814
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/279,832
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,833
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/279,863
; PRIOR FILING DATE: 2001-03-29
; PRIOR APPLICATION NUMBER: 60/283,889
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,447
; PRIOR FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: 60/286,683
; PRIOR FILING DATE: 2001-04-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 411
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 102
; LENGTH: 312
; TYPE: PRT
; ORGANISM: Bos taurus
US-10-038-854-102

Query Match      0.8%; Score 9; DB 16; Length 312;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVLA 1162
      |||||
Db 3 LLLALLVLA 11

RESULT 63
US-10-369-493-2470
; Sequence 2470, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 2470
; LENGTH: 371
; TYPE: PRT
; ORGANISM: Schizosaccharomyces pombe
US-10-369-493-2470

Query Match      0.8%; Score 9; DB 15; Length 371;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 378 DGVLLGAVG 386
      |||||
```

```
Db 69 DGVLLGAVG 77

RESULT 64
US-09-350-259-20
; Sequence 20, Application US/09350259
; Patent No. US20020062008A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20020062008A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
US-09-350-259-20

Query Match      0.7%; Score 8; DB 9; Length 17;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
      |||||
Db 1 LVVGAPLE 8

RESULT 65
US-09-891-943-20
; Sequence 20, Application US/09891943
; Publication No. US20030077278A1
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. US20030077278A1el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/891,943
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
US-09-891-943-20

Query Match      0.7%; Score 8; DB 10; Length 17;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
```

```
Db      1 LVVGAPLE 8
|||||||
RESULT 66
US-10-144-259-27
; Sequence 27, Application US/10144259
; Publication No. US20030109691A1
; GENERAL INFORMATION:
; APPLICANT: Arnaout, M. Amin
; APPLICANT: Li, Rui
; APPLICANT: Xiong, Jian-Ping
; TITLE OF INVENTION: VARIANT INTEGRIN POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: 00786-548001
; CURRENT APPLICATION NUMBER: US/10/144,259
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: US 03/758,493
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: US 60/221,950
; PRIOR FILING DATE: 2000-07-31
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-144-259-27

Query Match      0.7%; Score 8; DB 14; Length 43;
Best Local Similarity 100.0%; Fred. No. 28;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      365 EMSQTGFS 372
|||||||
Db      32 EMSQTGFS 39

RESULT 67
US-09-978-295A-85
; Sequence 85, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Destrochers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Cao, Wei-Quang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC11
; CURRENT APPLICATION NUMBER: US/09/978,295A
```

;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083332
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339

;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
Db 12 GLLLLALL 19

RESULT 68

US-09-978-697-85
; Sequence 85, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Shezuan
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin J.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas P.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03

1, PRIOR APPLICATION NUMBER: 60/065311
2, PRIOR FILING DATE: 1997-11-13
3, PRIOR APPLICATION NUMBER: 60/066364
4, PRIOR FILING DATE: 1997-11-21
5, PRIOR APPLICATION NUMBER: 60/077450
6, PRIOR FILING DATE: 1998-03-10
7, PRIOR APPLICATION NUMBER: 60/077632
8, PRIOR FILING DATE: 1998-03-11
9, PRIOR APPLICATION NUMBER: 60/077641
10, PRIOR FILING DATE: 1998-03-11
11, PRIOR APPLICATION NUMBER: 60/077649
12, PRIOR FILING DATE: 1998-03-11
13, PRIOR APPLICATION NUMBER: 60/077791
14, PRIOR FILING DATE: 1998-03-12
15, PRIOR APPLICATION NUMBER: 60/078004
16, PRIOR FILING DATE: 1998-03-13
17, PRIOR APPLICATION NUMBER: 60/078886
18, PRIOR FILING DATE: 1998-03-20
19, PRIOR APPLICATION NUMBER: 60/078936
20, PRIOR FILING DATE: 1998-03-20
21, PRIOR APPLICATION NUMBER: 60/078910
22, PRIOR FILING DATE: 1998-03-20
23, PRIOR APPLICATION NUMBER: 60/078939
24, PRIOR FILING DATE: 1998-03-20
25, PRIOR APPLICATION NUMBER: 60/079294
26, PRIOR FILING DATE: 1998-03-25
27, PRIOR APPLICATION NUMBER: 60/079656
28, PRIOR FILING DATE: 1998-03-26
29, PRIOR APPLICATION NUMBER: 60/079664
30, PRIOR FILING DATE: 1998-03-27
31, PRIOR APPLICATION NUMBER: 60/079689
32, PRIOR FILING DATE: 1998-03-27
33, PRIOR APPLICATION NUMBER: 60/079663
34, PRIOR FILING DATE: 1998-03-27
35, PRIOR APPLICATION NUMBER: 60/079728
36, PRIOR FILING DATE: 1998-03-27
37, PRIOR APPLICATION NUMBER: 60/079786
38, PRIOR FILING DATE: 1998-03-27
39, PRIOR APPLICATION NUMBER: 60/079920
40, PRIOR FILING DATE: 1998-03-30
41, PRIOR APPLICATION NUMBER: 60/079923
42, PRIOR FILING DATE: 1998-03-30
43, PRIOR APPLICATION NUMBER: 60/080105
44, PRIOR FILING DATE: 1998-03-31
45, PRIOR APPLICATION NUMBER: 60/080107
46, PRIOR FILING DATE: 1998-03-31
47, PRIOR APPLICATION NUMBER: 60/080165
48, PRIOR FILING DATE: 1998-03-31
49, PRIOR APPLICATION NUMBER: 60/080194
50, PRIOR FILING DATE: 1998-03-31
51, PRIOR APPLICATION NUMBER: 60/080327
52, PRIOR FILING DATE: 1998-04-01
53, PRIOR APPLICATION NUMBER: 60/080328
54, PRIOR FILING DATE: 1998-04-01
55, PRIOR APPLICATION NUMBER: 60/080333
56, PRIOR FILING DATE: 1998-04-01
57, PRIOR APPLICATION NUMBER: 60/080334
58, PRIOR FILING DATE: 1998-04-01
59, PRIOR APPLICATION NUMBER: 60/081070
60, PRIOR FILING DATE: 1998-04-08
61, PRIOR APPLICATION NUMBER: 60/081049
62, PRIOR FILING DATE: 1998-04-08
63, PRIOR APPLICATION NUMBER: 60/081071
64, PRIOR FILING DATE: 1998-04-08
65, PRIOR APPLICATION NUMBER: 60/081195
66, PRIOR FILING DATE: 1998-04-08
67, PRIOR APPLICATION NUMBER: 60/081203
68, PRIOR FILING DATE: 1998-04-09
69, PRIOR APPLICATION NUMBER: 60/081229
70, PRIOR FILING DATE: 1998-04-09
71, PRIOR APPLICATION NUMBER: 60/081955
72, PRIOR FILING DATE: 1998-04-15
73, PRIOR APPLICATION NUMBER: 60/081817


```
/ PRIOR APPLICATION NUMBER: 60/085700
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085689
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085579
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085580
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085573
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085704
/ PRIOR FILING DATE: 1998-05-15
/ PRIOR APPLICATION NUMBER: 60/085597

Query Match      0.7%; Score 8; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1152 GULLLALL 1159
Db      12 GULLLALL 19

RESULT 69
US-09-978-192A-85
/ Sequence 85, Application US/09978192A
/ Patent No. US20020177553A1
/ GENERAL INFORMATION:
/ APPLICANT: Ashkenazi, Avi
/ APPLICANT: Baker Kevin P.
/ APPLICANT: Botstein, David
/ APPLICANT: Desnovers, Luc
/ APPLICANT: Eaton, Dan
/ APPLICANT: Ferrara, Napoleon
/ APPLICANT: Filvaroff, Ellen
/ APPLICANT: Fong, Sherman
/ APPLICANT: Gac, Wei-Qiang
/ APPLICANT: Gerber, Hanspeter
/ APPLICANT: Gerritsen, Mary E.
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Hillan, Kenneth J.
/ APPLICANT: Kljavin, Ivar J.
/ APPLICANT: Kuo, Sophia S.
/ APPLICANT: Napier, Mary A.
/ APPLICANT: Pan, James;
/ APPLICANT: Paoni, Nicholas F.
/ APPLICANT: Roy, Margaret Ann
/ APPLICANT: Shelton, David L.
/ APPLICANT: Stewart, Timothy A.
/ APPLICANT: Tumas, Daniel
/ APPLICANT: Williams, P. Mickey
/ APPLICANT: Wood, William I.
/ TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
/ TITLE OF INVENTION: Acids Encoding the Same
/ FILE REFERENCE: P2630P1C9
/ CURRENT APPLICATION NUMBER: US/09/978,192A
/ CURRENT FILING DATE: 2001-10-15
/ PRIOR APPLICATION NUMBER: 09/918585
/ PRIOR FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: 60/062250
/ PRIOR FILING DATE: 1997-10-17
/ PRIOR APPLICATION NUMBER: 60/064249
/ PRIOR FILING DATE: 1997-11-03
/ PRIOR APPLICATION NUMBER: 60/065311
/ PRIOR FILING DATE: 1997-11-13
/ PRIOR APPLICATION NUMBER: 60/066364
/ PRIOR FILING DATE: 1997-11-21
/ PRIOR APPLICATION NUMBER: 60/077450
/ PRIOR FILING DATE: 1998-03-10
/ PRIOR APPLICATION NUMBER: 60/077632
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077641
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077649
/ PRIOR FILING DATE: 1998-03-11
/ PRIOR APPLICATION NUMBER: 60/077791
/ PRIOR FILING DATE: 1998-03-12
/ PRIOR APPLICATION NUMBER: 60/078004
/ PRIOR FILING DATE: 1998-03-13
/ PRIOR APPLICATION NUMBER: 60/078886
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078936
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078910
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/078939
/ PRIOR FILING DATE: 1998-03-20
/ PRIOR APPLICATION NUMBER: 60/079294
/ PRIOR FILING DATE: 1998-03-25
/ PRIOR APPLICATION NUMBER: 60/079656
/ PRIOR FILING DATE: 1998-03-26
/ PRIOR APPLICATION NUMBER: 60/079664
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079689
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079663
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079728
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079786
/ PRIOR FILING DATE: 1998-03-27
/ PRIOR APPLICATION NUMBER: 60/079920
/ PRIOR FILING DATE: 1998-03-30
/ PRIOR APPLICATION NUMBER: 60/079923
/ PRIOR FILING DATE: 1998-03-30
/ PRIOR APPLICATION NUMBER: 60/080105
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080107
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080165
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080194
/ PRIOR FILING DATE: 1998-03-31
/ PRIOR APPLICATION NUMBER: 60/080327
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/080328
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/080333
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/080334
/ PRIOR FILING DATE: 1998-04-01
/ PRIOR APPLICATION NUMBER: 60/081070
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081049
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081071
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081195
/ PRIOR FILING DATE: 1998-04-08
/ PRIOR APPLICATION NUMBER: 60/081203
/ PRIOR FILING DATE: 1998-04-09
/ PRIOR APPLICATION NUMBER: 60/081229
/ PRIOR FILING DATE: 1998-04-09
/ PRIOR APPLICATION NUMBER: 60/081955
/ PRIOR FILING DATE: 1998-04-15
/ PRIOR APPLICATION NUMBER: 60/081817
/ PRIOR FILING DATE: 1998-04-15
/ PRIOR APPLICATION NUMBER: 60/081819
/ PRIOR FILING DATE: 1998-04-15
/ PRIOR APPLICATION NUMBER: 60/081952
/ PRIOR FILING DATE: 1998-04-15
/ PRIOR APPLICATION NUMBER: 60/081838
/ PRIOR FILING DATE: 1998-04-15
```

; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580

; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 9; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALL 1159

Db 12 GLLALL 19
|||||||

RESULT 70

US-09-999-832A-85
; Sequence 85, Application US/0999832A

; Publication No. US20020192706A1

; GENERAL INFORMATION:

; APPLICANT: Ashkenazi, Avi

; APPLICANT: Baker Kevin P.

; APPLICANT: Botstein, David

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan

; APPLICANT: Ferrara, Napoleon

; APPLICANT: Filvaroff, Ellen

; APPLICANT: Fong, Sherman

; APPLICANT: Gao, Wei-Qiang

; APPLICANT: Gerber, Hanspeter

; APPLICANT: Gerritsen, Mary E.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Grimaldi, J. Christopher

; APPLICANT: Gurney, Austin L.

; APPLICANT: Hillan, Kenneth J.

; APPLICANT: Kljavin, Ivar J.

; APPLICANT: Kuo, Sophia S.

; APPLICANT: Napier, Mary A.

; APPLICANT: Pan, James;

; APPLICANT: Paoni, Nicholas F.

; APPLICANT: Roy, Margaret Ann

; APPLICANT: Shelton, David L.

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Tumas, Daniel

; APPLICANT: Williams, P. Mickey

; APPLICANT: Wood, William I.

; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic

; FILE REFERENCE: P2630P1C63

; CURRENT APPLICATION NUMBER: US/09/999,832A

; CURRENT FILING DATE: 2001-10-24

; PRIOR APPLICATION NUMBER: 09/918585

; PRIOR FILING DATE: 2001-07-30

; PRIOR APPLICATION NUMBER: 60/062250

; PRIOR FILING DATE: 1997-10-17

; PRIOR APPLICATION NUMBER: 60/064249

; PRIOR FILING DATE: 1997-11-03

; PRIOR APPLICATION NUMBER: 60/065311

; PRIOR FILING DATE: 1997-11-13

; PRIOR APPLICATION NUMBER: 60/066364

; PRIOR FILING DATE: 1997-11-21

; PRIOR APPLICATION NUMBER: 60/077450

; PRIOR FILING DATE: 1998-03-10

; PRIOR APPLICATION NUMBER: 60/077632

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077641

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077649

; PRIOR FILING DATE: 1998-03-11

; PRIOR APPLICATION NUMBER: 60/077791

; PRIOR FILING DATE: 1998-03-12

1	PRIOR FILING DATE: 1998-04-22	
2	PRIOR APPLICATION NUMBER: 60/082700	
3	PRIOR FILING DATE: 1998-04-22	
4	PRIOR APPLICATION NUMBER: 60/082797	
5	PRIOR FILING DATE: 1998-04-22	
6	PRIOR APPLICATION NUMBER: 60/082796	
7	PRIOR FILING DATE: 1998-04-23	
8	PRIOR APPLICATION NUMBER: 60/083336	
9	PRIOR FILING DATE: 1998-04-27	
10	PRIOR APPLICATION NUMBER: 60/083322	
11	PRIOR FILING DATE: 1998-04-28	
12	PRIOR APPLICATION NUMBER: 60/083392	
13	PRIOR FILING DATE: 1998-04-29	
14	PRIOR APPLICATION NUMBER: 60/083495	
15	PRIOR FILING DATE: 1998-04-29	
16	PRIOR APPLICATION NUMBER: 60/083496	
17	PRIOR FILING DATE: 1998-04-29	
18	PRIOR APPLICATION NUMBER: 60/083499	
19	PRIOR FILING DATE: 1998-04-29	
20	PRIOR APPLICATION NUMBER: 60/083545	
21	PRIOR FILING DATE: 1998-04-29	
22	PRIOR APPLICATION NUMBER: 60/083554	
23	PRIOR FILING DATE: 1998-04-29	
24	PRIOR APPLICATION NUMBER: 60/083558	
25	PRIOR FILING DATE: 1998-04-29	
26	PRIOR APPLICATION NUMBER: 60/083559	
27	PRIOR FILING DATE: 1998-04-29	
28	PRIOR APPLICATION NUMBER: 60/083500	
29	PRIOR FILING DATE: 1998-04-29	
30	PRIOR APPLICATION NUMBER: 60/083742	
31	PRIOR FILING DATE: 1998-04-30	
32	PRIOR APPLICATION NUMBER: 60/084366	
33	PRIOR FILING DATE: 1998-05-05	
34	PRIOR APPLICATION NUMBER: 60/084414	
35	PRIOR FILING DATE: 1998-05-06	
36	PRIOR APPLICATION NUMBER: 60/084441	
37	PRIOR FILING DATE: 1998-05-06	
38	PRIOR APPLICATION NUMBER: 60/084637	
39	PRIOR FILING DATE: 1998-05-07	
40	PRIOR APPLICATION NUMBER: 60/084639	
41	PRIOR FILING DATE: 1998-05-07	
42	PRIOR APPLICATION NUMBER: 60/084640	
43	PRIOR FILING DATE: 1998-05-07	
44	PRIOR APPLICATION NUMBER: 60/084598	
45	PRIOR FILING DATE: 1998-05-07	
46	PRIOR APPLICATION NUMBER: 60/084600	
47	PRIOR FILING DATE: 1998-05-07	
48	PRIOR APPLICATION NUMBER: 60/084627	
49	PRIOR FILING DATE: 1998-05-07	
50	PRIOR APPLICATION NUMBER: 60/084643	
51	PRIOR FILING DATE: 1998-05-07	
52	PRIOR APPLICATION NUMBER: 60/085339	
53	PRIOR FILING DATE: 1998-05-13	
54	PRIOR APPLICATION NUMBER: 60/085338	
55	PRIOR FILING DATE: 1998-05-13	
56	PRIOR APPLICATION NUMBER: 60/085323	
57	PRIOR FILING DATE: 1998-05-13	
58	PRIOR APPLICATION NUMBER: 60/085362	
59	PRIOR FILING DATE: 1998-05-15	
60	PRIOR APPLICATION NUMBER: 60/085700	
61	PRIOR FILING DATE: 1998-05-15	
62	PRIOR APPLICATION NUMBER: 60/085669	
63	PRIOR FILING DATE: 1998-05-15	
64	PRIOR APPLICATION NUMBER: 60/085579	
65	PRIOR FILING DATE: 1998-05-15	
66	PRIOR APPLICATION NUMBER: 60/085560	
67	PRIOR FILING DATE: 1998-05-15	
68	PRIOR APPLICATION NUMBER: 60/085573	
69	PRIOR FILING DATE: 1998-05-15	
70	PRIOR APPLICATION NUMBER: 60/085704	
71	PRIOR FILING DATE: 1998-05-15	
72	PRIOR APPLICATION NUMBER: 60/085697	

Query Match 0.7% Score 8; DB 9; Length 67;
 Best Local Similarity 100.0%; Pred. No. 41;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
 Db 12 GLLLLALL 19

RESULT 71

US-09-978-189-85
 ; Sequence 85, Application US/09978189
 ; Publication No. US2003004102A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ashkenazi, Avi
 ; APPLICANT: Baker Kevin P.
 ; APPLICANT: Botstein, David
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan
 ; APPLICANT: Ferrara, Napoleon
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Fong, Sherman
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerber, Hanspeter
 ; APPLICANT: Gottschalk, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, J. Christopher
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Hillan, Kenneth J.
 ; APPLICANT: Kijavini, Ivar J.
 ; APPLICANT: Kuo, Sophia S.
 ; APPLICANT: Napier, Mary A.
 ; APPLICANT: Pan, James;
 ; APPLICANT: Paoni, Nicholas F.
 ; APPLICANT: Roy, Margaret Ann
 ; APPLICANT: Shelton, David L.
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Williams, P. Mickey
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
 ; TITLE OF INVENTION: Acids Encoding the Same
 ; FILE REFERENCE: P2630P1C7
 ; CURRENT APPLICATION NUMBER: US/09/978,189
 ; CURRENT FILING DATE: 2001-10-15
 ; PRIOR APPLICATION NUMBER: 09/918585
 ; PRIOR FILING DATE: 2001-07-30
 ; PRIOR APPLICATION NUMBER: 60/062250
 ; PRIOR FILING DATE: 1997-10-17
 ; PRIOR APPLICATION NUMBER: 60/064249
 ; PRIOR FILING DATE: 1997-11-03
 ; PRIOR APPLICATION NUMBER: 60/065311
 ; PRIOR FILING DATE: 1997-11-13
 ; PRIOR APPLICATION NUMBER: 60/066364
 ; PRIOR FILING DATE: 1997-11-21
 ; PRIOR APPLICATION NUMBER: 60/077450
 ; PRIOR FILING DATE: 1998-03-10
 ; PRIOR APPLICATION NUMBER: 60/077632
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077641
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077649
 ; PRIOR FILING DATE: 1998-03-11
 ; PRIOR APPLICATION NUMBER: 60/077791
 ; PRIOR FILING DATE: 1998-03-12
 ; PRIOR APPLICATION NUMBER: 60/078004
 ; PRIOR FILING DATE: 1998-03-13
 ; PRIOR APPLICATION NUMBER: 60/078886
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078936
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078910

; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/078939
 ; PRIOR FILING DATE: 1998-03-20
 ; PRIOR APPLICATION NUMBER: 60/079294
 ; PRIOR FILING DATE: 1998-03-25
 ; PRIOR APPLICATION NUMBER: 60/079656
 ; PRIOR FILING DATE: 1998-03-26
 ; PRIOR APPLICATION NUMBER: 60/079664
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079689
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079663
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079728
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079786
 ; PRIOR FILING DATE: 1998-03-27
 ; PRIOR APPLICATION NUMBER: 60/079920
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/079923
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 60/080105
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080107
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080165
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080194
 ; PRIOR FILING DATE: 1998-03-31
 ; PRIOR APPLICATION NUMBER: 60/080327
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080328
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080333
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/080334
 ; PRIOR FILING DATE: 1998-04-01
 ; PRIOR APPLICATION NUMBER: 60/081070
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081049
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081071
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081195
 ; PRIOR FILING DATE: 1998-04-08
 ; PRIOR APPLICATION NUMBER: 60/081203
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081229
 ; PRIOR FILING DATE: 1998-04-09
 ; PRIOR APPLICATION NUMBER: 60/081955
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081817
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081819
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081952
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/081838
 ; PRIOR FILING DATE: 1998-04-15
 ; PRIOR APPLICATION NUMBER: 60/082568
 ; PRIOR FILING DATE: 1998-04-21
 ; PRIOR APPLICATION NUMBER: 60/082569
 ; PRIOR FILING DATE: 1998-04-21
 ; PRIOR APPLICATION NUMBER: 60/082704
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082804
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082700
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082797
 ; PRIOR FILING DATE: 1998-04-22
 ; PRIOR APPLICATION NUMBER: 60/082796
 ; PRIOR FILING DATE: 1998-04-23

; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085592
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085699
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
| | | | |
Db 12 GLLLLALL 19

RESULT 72

US-09-978-608A-85
; Sequence 85, Application US/09978608A
; Publication No. US20030045462A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kllavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C22
; CURRENT APPLICATION NUMBER: US/09/978,608A
; CURRENT FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 85
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-608A-85

Query Match 0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1152 GLLLLALL 1159
| | | | |
Db 12 GLLLLALL 19

RESULT 73

US-09-978-585A-85
; Sequence 85, Application US/09978585A
; Publication No. US20030049633A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey

```

; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James.
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC15
; CURRENT APPLICATION NUMBER: US/09/978,585A
; CURRENT FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 85
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-978-585A-85

Query Match 0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19
|||||
|||||

RESULT 74
US-09-978-191A-85
; Sequence 85, Application US/09978191A
; Publication No. US20030050239A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kijavini, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC4
; CURRENT APPLICATION NUMBER: US/09/978,191A
; CURRENT FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08

```

;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339

;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;

Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
Db 12 GLLLLALL 19
|||||

RESULT 75

US-09-978-403A-85
; Sequence 85, Application US/09978403A
; Publication No. US20030050240A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kllavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE REFERENCE: P2630PIC17
; CURRENT APPLICATION NUMBER: US/09/978,403A
; CURRENT FILING DATE: 2002-03-19
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	524	525
---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Query Match 0.7%; Score 8; DB 10; Length 67;
Best Local Similarity 100.0%; Pred. No. 41;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLLALL 1159
|||
Db 12 GLLLLALL 19

Search completed: September 21, 2004, 14:09:50
Job time : 85 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 21, 2004, 13:29:15 ; Search time 22 Seconds

(without alignments)
2787.602 Million cell updates/sec

Title: US-09-980-403-2

Perfect score: 1188

Sequence: 1 MDLPRGLVVAWALSLWPGFT.....FRSARRRPGLDPTPKYLE 1188

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 389414 seqs, 51625971 residues

Word size : 0

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Database : Issued Patents AA.*

1: /cgn2_6/ptodata/2/iaa/5A-COMB.pep.*
2: /cgn2_6/ptodata/2/iaa/5B-COMB.pep.*
3: /cgn2_6/ptodata/2/iaa/6A-COMB.pep.*
4: /cgn2_6/ptodata/2/iaa/6B-COMB.pep.*
5: /cgn2_6/ptodata/2/iaa/PCTUS-COMB.pep.*
6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	1.8	158	4	US-09-312-283C-391
2	17	1.4	85	4	US-09-312-283C-390
3	15	1.3	214	4	US-10-061-658-5
4	15	1.3	214	4	US-10-061-658-6
5	15	1.3	214	4	US-10-061-658-9
6	14	1.2	1183	4	US-09-532-310B-5
7	11	0.9	22	6	5196511-9
8	11	0.9	1183	4	US-09-532-310B-6
9	9	0.8	9	1	US-08-462-128-34
10	9	0.8	9	1	US-08-463-180-34
11	8	0.7	17	1	US-08-173-497-20
12	8	0.7	17	1	US-08-286-889-20
13	8	0.7	17	1	US-08-485-618-20
14	8	0.7	17	1	US-08-362-652-20
15	8	0.7	17	2	US-08-605-672-20
16	8	0.7	17	2	US-08-482-293A-20
17	8	0.7	17	2	US-08-943-363-20
18	8	0.7	17	3	US-09-193-043-20
19	8	0.7	17	4	US-09-688-307A-20
20	8	0.7	17	4	US-09-350-259-20
21	8	0.7	76	1	US-08-848-252-4
22	8	0.7	76	3	US-09-083-521-5
23	8	0.7	80	1	US-08-848-252-2
24	8	0.7	95	1	US-07-626-618A-2
25	8	0.7	95	1	US-07-928-611-2
26	8	0.7	95	1	US-08-333-977-2
27	8	0.7	95	2	US-08-487-811A-2

28	8	0.7	95	3	US-09-060-694-2	Sequence 2, Appli
29	8	0.7	95	4	US-09-378-074-2	Sequence 2, Appli
30	8	0.7	95	5	PCT-US93-07370-2	Sequence 2, Appli
31	8	0.7	97	4	US-08-973-544-4	Sequence 4, Appli
32	8	0.7	104	4	US-08-973-544-3	Sequence 3, Appli
33	8	0.7	112	4	US-09-489-039A-10236	Sequence 10236, A
34	8	0.7	179	4	US-09-419-568B-28	Sequence 28, Appli
35	8	0.7	179	4	US-09-354-243B-28	Sequence 28, Appli
36	8	0.7	179	4	US-09-870-574-2	Sequence 2, Appli
37	8	0.7	183	4	US-09-252-991A-25905	Sequence 25905, A
38	8	0.7	234	3	US-09-040-483-5	Sequence 5, Appli
39	8	0.7	234	4	US-08-740-036-5	Sequence 5, Appli
40	8	0.7	315	1	US-08-118-270-28	Sequence 28, Appli
41	8	0.7	315	5	PCT-US93-08528-28	Sequence 28, Appli
42	8	0.7	345	4	US-09-489-039A-7417	Sequence 7417, Ap
43	8	0.7	374	3	US-09-046-736-4	Sequence 4, Appli
44	8	0.7	387	1	US-07-626-618A-17	Sequence 17, Appli
45	8	0.7	387	1	US-08-086-439C-3	Sequence 3, Appli
46	8	0.7	387	1	US-08-056-051-2	Sequence 2, Appli
47	8	0.7	387	1	US-07-928-611-18	Sequence 18, Appli
48	8	0.7	387	1	US-08-333-977-17	Sequence 17, Appli
49	8	0.7	387	1	US-08-449-153-1	Sequence 1, Appli
50	8	0.7	387	1	US-08-434-877-3	Sequence 3, Appli
51	8	0.7	387	1	US-08-449-651-1	Sequence 1, Appli
52	8	0.7	387	3	US-08-487-811A-18	Sequence 18, Appli
53	8	0.7	387	3	US-08-833-703-1	Sequence 1, Appli
54	8	0.7	387	3	US-08-475-742-4	Sequence 4, Appli
55	8	0.7	387	3	US-09-060-694-18	Sequence 18, Appli
56	8	0.7	387	4	US-09-378-074-18	Sequence 18, Appli
57	8	0.7	387	4	US-08-261-293-4	Sequence 4, Appli
58	8	0.7	387	5	PCT-US93-07370-18	Sequence 18, Appli
59	8	0.7	419	1	US-08-056-051-4	Sequence 4, Appli
60	8	0.7	419	1	US-07-928-611-20	Sequence 20, Appli
61	8	0.7	419	2	US-08-487-811A-20	Sequence 20, Appli
62	8	0.7	419	3	US-09-060-694-20	Sequence 20, Appli
63	8	0.7	419	4	US-09-378-074-20	Sequence 20, Appli
64	8	0.7	419	5	PCT-US93-07370-20	Sequence 20, Appli
65	8	0.7	453	4	US-09-252-991A-17394	Sequence 17394, A
66	8	0.7	467	1	US-08-056-051-6	Sequence 6, Appli
67	8	0.7	467	1	US-07-928-611-22	Sequence 22, Appli
68	8	0.7	467	2	US-08-487-811A-22	Sequence 22, Appli
69	8	0.7	467	3	US-09-046-736-2	Sequence 2, Appli
70	8	0.7	467	3	US-09-060-694-22	Sequence 22, Appli
71	8	0.7	467	4	US-09-378-074-22	Sequence 22, Appli
72	8	0.7	467	5	PCT-US93-07370-22	Sequence 22, Appli
73	8	0.7	494	1	US-08-485-618-103	Sequence 103, App
74	8	0.7	494	2	US-08-605-672-103	Sequence 103, App
75	8	0.7	494	2	US-08-482-293A-103	Sequence 103, App
76	8	0.7	494	2	US-08-943-363-103	Sequence 103, App
77	8	0.7	494	3	US-09-193-043-103	Sequence 103, App
78	8	0.7	494	4	US-09-688-307A-103	Sequence 103, App
79	8	0.7	494	4	US-09-350-259-103	Sequence 103, App
80	8	0.7	605	2	US-08-472-666-1	Sequence 1, Appli
81	8	0.7	605	5	PCT-US96-07615-1	Sequence 1, Appli
82	8	0.7	628	4	US-09-252-991A-32727	Sequence 32727, A
83	8	0.7	628	4	US-09-352-991A-25547	Sequence 25547, A
84	8	0.7	1151	1	US-08-286-889-37	Sequence 37, Appli
85	8	0.7	1151	1	US-08-485-618-37	Sequence 37, Appli
86	8	0.7	1151	1	US-08-362-652-37	Sequence 37, Appli
87	8	0.7	1151	2	US-08-605-672-37	Sequence 37, Appli
88	8	0.7	1151	2	US-08-482-293A-37	Sequence 37, Appli
89	8	0.7	1151	2	US-08-943-363-37	Sequence 37, Appli
90	8	0.7	1151	3	US-09-193-043-37	Sequence 37, Appli
91	8	0.7	1151	4	US-09-688-307A-37	Sequence 37, Appli
92	8	0.7	1151	4	US-09-350-259-37	Sequence 37, Appli
93	8	0.7	1152	2	US-08-476-062A-43	Sequence 43, Appli
94	8	0.7	1152	5	PCT-US96-01314-43	Sequence 43, Appli
95	8	0.7	1152	6	5424399-2	Patent No. 5424399
96	8	0.7	1153	1	US-08-173-497-3	Sequence 3, Appli
97	8	0.7	1153	1	US-08-286-889-3	Sequence 3, Appli
98	8	0.7	1153	1	US-08-485-618-3	Sequence 3, Appli
99	8	0.7	1153	1	US-08-362-652-3	Sequence 3, Appli
100	8	0.7	1153	2	US-08-605-672-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1

US-09-312-283C-391
; Sequence 391, Application US/09312283C
; Patent No. 6573095

; GENERAL INFORMATION:
; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated from Skin Cells

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312,283C

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 391

; LENGTH: 158

; TYPE: PRT

; ORGANISM: Mouse

US-09-312-283C-391

Query Match

Best Local Similarity 1.8%; Score 21; DB 4; Length 158;

Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

636 RPVVQINASHPEPSKINIFH 656

Db

57 RPVVQINASHPEPSKINIFH 77

RESULT 2

US-09-312-283C-390
; Sequence 390, Application US/09312283C
; Patent No. 6573095

; GENERAL INFORMATION:
; APPLICANT: Watson, James D.

; APPLICANT: Strachan, Lorna

; APPLICANT: Sleeman, Matthew

; APPLICANT: Onrust, Rene

; APPLICANT: Murison, James G.

; APPLICANT: Kumble, Krishanand D.

; TITLE OF INVENTION: Compositions Isolated from Skin Cells

; FILE REFERENCE: 11000.1011c2

; CURRENT APPLICATION NUMBER: US/09/312,283C

; CURRENT FILING DATE: 1999-05-14

; NUMBER OF SEQ ID NOS: 425

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 390

; LENGTH: 85

; TYPE: PRT

; ORGANISM: Mouse

US-09-312-283C-390

Query Match

Best Local Similarity 1.4%; Score 17; DB 4; Length 85;

Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

528 NGTLKDSHSYQNAFSG 544

Db

28 NGTLKDSHSYQNAFSG 44

RESULT 3

US-10-061-658-5

; Sequence 5, Application US/10061658
; Patent No. 6652856
; GENERAL INFORMATION:

; APPLICANT: Biogen, Inc.

; APPLICANT: Gotwals, Philip

; TITLE OF INVENTION: Method for the Treatment of Fibrosis

; FILE REFERENCE: A073US

; CURRENT APPLICATION NUMBER: US/10/061,658

; CURRENT FILING DATE: 2002-02-01

; PRIOR APPLICATION NUMBER: 60/137,214

; PRIOR FILING DATE: 1999-06-01

; PRIOR APPLICATION NUMBER: 60/130,847

; PRIOR FILING DATE: 1999-04-22

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 5

; LENGTH: 214

; TYPE: PRT

; ORGANISM: Rat

US-10-061-658-5

Query Match

Best Local Similarity 1.3%; Score 15; DB 4; Length 214;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

164 DIVVLGDGNSIYPW 178

Db

22 DIVVLGDGNSIYPW 36

RESULT 4

US-10-061-658-6
; Sequence 6, Application US/10061658
; Patent No. 6652856

; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.

; APPLICANT: Gotwals, Philip

; TITLE OF INVENTION: Method for the Treatment of Fibrosis

; FILE REFERENCE: A073US

; CURRENT APPLICATION NUMBER: US/10/061,658

; CURRENT FILING DATE: 2002-02-01

; PRIOR APPLICATION NUMBER: 60/137,214

; PRIOR FILING DATE: 1999-06-01

; PRIOR APPLICATION NUMBER: 60/130,847

; PRIOR FILING DATE: 1999-04-22

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 6

; LENGTH: 214

; TYPE: PRT

; ORGANISM: Homo sapien

US-10-061-658-6

Query Match

Best Local Similarity 1.3%; Score 15; DB 4; Length 214;

Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy

164 DIVVLGDGNSIYPW 178

Db

22 DIVVLGDGNSIYPW 36

RESULT 5

US-10-061-658-9
; Sequence 9, Application US/10061658
; Patent No. 6652856

; GENERAL INFORMATION:
; APPLICANT: Biogen, Inc.

; APPLICANT: Gotwals, Philip

; TITLE OF INVENTION: Method for the Treatment of Fibrosis

```
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-061-658-9

Query Match      1.3%; Score 15; DB 4; Length 214;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYP 178
      |||||||
Db      22 DIVVLGDSNSIYP 36

RESULT 6
US-09-532-310B-5
; Sequence 5, Application US/09532310B
; Patent No. 6596276
; GENERAL INFORMATION:
; APPLICANT: Senger, Donald R
;           Detmar, Michael
;           Claffey, Kevin P
; TITLE OF INVENTION: Method for inhibiting tumor
;                   angiogenesis in a living subject
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David Prashker, Esq.
; STREET: P.O. Box 5387
; CITY: Magnolia
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 01930
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage
; COMPUTER: Dell PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Microsoft Word version 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/532,310B
; FILING DATE: 22-Mar-2000
; CLASSIFICATION: Unknown
; ATTORNEY/AGENT INFORMATION:
; NAME: David Prashker, Esq.
; REGISTRATION NUMBER: 29,693
; REFERENCE/DOCKET NUMBER: BIS-036
; TELEPHONE: (978) 525-3794
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 1183 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:
US-09-532-310B-5

Query Match      1.2%; Score 14; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYP 177
      |||||||
; FILE REFERENCE: A073US
; CURRENT APPLICATION NUMBER: US/10/061,658
; CURRENT FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/137,214
; PRIOR FILING DATE: 1999-06-01
; PRIOR APPLICATION NUMBER: 60/130,847
; PRIOR FILING DATE: 1999-04-22
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 214
; TYPE: PRT
; ORGANISM: Homo sapien
US-10-061-658-9

Query Match      1.3%; Score 15; DB 4; Length 214;
Best Local Similarity 100.0%; Pred. No. 6.6e-06;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      164 DIVVLGDSNSIYP 178
      |||||||
Db      22 DIVVLGDSNSIYP 36

RESULT 7
US-09-532-310B-6
; Sequence 5, Application US/09532310B
; Patent No. 6596276
; GENERAL INFORMATION:
; APPLICANT: Senger, Donald R
;           Detmar, Michael
;           Claffey, Kevin P
; TITLE OF INVENTION: Method for inhibiting tumor
;                   angiogenesis in a living subject
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David Prashker, Esq.
; STREET: P.O. Box 5387
; CITY: Magnolia
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 01930
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.40 Mb storage
; COMPUTER: Dell PC
; OPERATING SYSTEM: MS DOS
; SOFTWARE: Microsoft Word version 97
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/532,310B
; FILING DATE: 22-Mar-2000
; CLASSIFICATION: Unknown
; ATTORNEY/AGENT INFORMATION:
; NAME: David Prashker, Esq.
; REGISTRATION NUMBER: 29,693
; REFERENCE/DOCKET NUMBER: BIS-036
; TELEPHONE: (978) 525-3794
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; TYPE: amino acid
; LENGTH: 1183 amino acids
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-532-310B-6

Query Match      0.9%; Score 11; DB 4; Length 1183;
Best Local Similarity 100.0%; Pred. No. 0.019;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 495 TDVLLVGAPMY 505
 |||||
 Db 506 TDVLLVGAPMY 516

RESULT 9
 US-08-462-128-34
 ; Sequence 34, Application US/08462128
 ; Patent No. 5686059
 ; GENERAL INFORMATION:
 ; APPLICANT: Goetinck, Paul F.
 ; APPLICANT: Tondravi, Mehرداد
 ; TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR
 ; TITLE OF INVENTION: USE
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lahive & Cockfield
 ; STREET: 60 State Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/462,128
 FILING DATE:
 CLASSIFICATION: 514
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/006,096
 FILING DATE: 15-JAN-1993
 APPLICATION NUMBER: US 07/866,403
 FILING DATE: 10-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Paul L.
 REGISTRATION NUMBER: 35,965
 REFERENCE/DOCKET NUMBER: MGP-005
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

US-08-462-128-34
 Query Match 0.8%; Score 9; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 164 DIVVLDS 172
 |||||
 Db 1 DIVVLDS 9

RESULT 10
 US-08-463-180-34
 ; Sequence 34, Application US/08463180
 ; Patent No. 5741670
 ; GENERAL INFORMATION:
 ; APPLICANT: Goetinck, Paul F.
 ; APPLICANT: Tondravi, Mehرداد
 ; TITLE OF INVENTION: CARTILAGE MATRIX PROTEIN AND METHODS FOR
 ; TITLE OF INVENTION: USE
 ; NUMBER OF SEQUENCES: 43
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lahive & Cockfield
 ; STREET: 60 State Street

CITY: Boston
 STATE: MA
 COUNTRY: USA
 ZIP: 02109
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/463,180
 FILING DATE:
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 08/006,096
 FILING DATE: 15-JAN-1993
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/866,403
 FILING DATE: 10-APR-1992
 ATTORNEY/AGENT INFORMATION:
 NAME: Myers, Paul L.
 REGISTRATION NUMBER: 35,965
 REFERENCE/DOCKET NUMBER: MGP-005DV
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 227-7400
 TELEFAX: (617) 227-5941
 INFORMATION FOR SEQ ID NO: 34:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 9 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

US-08-463-180-34
 Query Match 0.8%; Score 9; DB 1; Length 9;
 Best Local Similarity 100.0%; Pred. No. 3e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 164 DIVVLDS 172
 |||||
 Db 1 DIVVLDS 9

RESULT 11
 US-08-173-497-20
 ; Sequence 20, Application US/08173497
 ; Patent No. 5437958
 ; GENERAL INFORMATION:
 ; APPLICANT: Gallatin, W. Michael
 ; APPLICANT: Van Der Vieren, Monica
 ; TITLE OF INVENTION: No. 5437958e1 Human 2 Integrin Alpha
 ; TITLE OF INVENTION: Subunit
 ; NUMBER OF SEQUENCES: 29
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
 ; STREET: 233 S. Wacker Drive, 6300 Sears Tower
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60606-6402
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/173,497
 FILING DATE:
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 5437958e1, Greta E.
 REGISTRATION NUMBER: 35,302
 REFERENCE/DOCKET NUMBER: 27866/31363

TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-173-497-20

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 12
US-08-286-889-20
Sequence 20, Application US/08286889
Patent No. 5470953
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Mich
TITLE OF INVENTION: No. 5470953el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/286,889
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32168
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-286-889-20

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 13
US-08-485-618-20
Sequence 20, Application US/08485618
Patent No. 5728533
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: No. 5728533el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-485-618-20

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 14
US-08-362-652-20
Sequence 20, Application US/08362652
Patent No. 5766850
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
TITLE OF INVENTION: No. 5766850el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 93
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/362,652
FILING DATE: 5-AUG-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32391
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-362-652-20

Query Match 0.7%; Score 8; DB 1; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 15
US-08-605-672-20
Sequence 20, Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889

FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:
LENGTH: 17 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-605-672-20

Query Match 0.7%; Score 8; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 16
US-08-482-293A-20
Sequence 20, Application US/08482293A
Patent No. 5831029
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/482,293A
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 20:
SEQUENCE CHARACTERISTICS:

;
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-482-293A-20

Query Match 0.7%; Score 8; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 17
US-08-943-363-20
; Sequence 20, Application US/08943363
; Patent No. 5837478
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; TITLE OF INVENTION: No. 5837478el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 114
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,363
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,497
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-943-363-20

Query Match 0.7%; Score 8; DB 2; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 18
US-09-193-043-20
; Sequence 20, Application US/09193043
; Patent No. 6251395
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 6251395el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/193,043
; CURRENT FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO: 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
US-09-193-043-20

Query Match 0.7%; Score 8; DB 3; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
Db 1 LVVGAPLE 8

RESULT 19
US-09-688-307A-20
; Sequence 20, Application US/09688307A
; Patent No. 6432404
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vlieten, Monica
; TITLE OF INVENTION: No. 6432404el Human Beta-2
; FILE REFERENCE: 27866/36646
; CURRENT APPLICATION NUMBER: US/09/688,307A
; CURRENT FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 09/193,043
; PRIOR FILING DATE: 1998-11-16
; PRIOR APPLICATION NUMBER: 08/605,672
; PRIOR FILING DATE: 1996-02-22
; PRIOR APPLICATION NUMBER: 08/173,497
; PRIOR FILING DATE: 1993-12-23
; PRIOR APPLICATION NUMBER: 08/286,889
; PRIOR FILING DATE: 1994-08-05
; PRIOR APPLICATION NUMBER: 08/362,652
; PRIOR FILING DATE: 1994-12-21
; PRIOR APPLICATION NUMBER: 08/943,363
; PRIOR FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO: 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 13
; OTHER INFORMATION: Xaa = any or unknown amino acid
US-09-688-307A-20

Query Match 0.7%; Score 8; DB 4; Length 17;

Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
|||||||
DB 1 LVVGAPLE 8

RESULT 20
US-09-350-259-20
; Sequence 20, Application US/09350259
; Patent No. 6620915
; GENERAL INFORMATION:
; APPLICANT: Gallatin, Michael W.
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 6620915el Human 2
; FILE REFERENCE: 27866/35004
; CURRENT APPLICATION NUMBER: US/09/350,259
; CURRENT FILING DATE: 1999-07-08
; EARLIER APPLICATION NUMBER: 09/193,043
; EARLIER FILING DATE: 1998-11-16
; EARLIER APPLICATION NUMBER: 08/173,497
; EARLIER FILING DATE: 1993-12-23
; EARLIER APPLICATION NUMBER: 08/286,889
; EARLIER FILING DATE: 1994-08-05
; EARLIER APPLICATION NUMBER: 08/362,652
; EARLIER FILING DATE: 1994-12-21
; EARLIER APPLICATION NUMBER: 08/943,363
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 17
; TYPE: PRT
; ORGANISM: dog
US-09-350-259-20

Query Match 0.7%; Score 8; DB 4; Length 17;
Best Local Similarity 100.0%; Pred. No. 3;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 56 LVVGAPLE 63
|||||||
DB 1 LVVGAPLE 8

RESULT 21
US-08-848-252-4
; Sequence 4, Application US/08848252
; Patent No. 5804177
; GENERAL INFORMATION:
; APPLICANT: Humphries, Keith R.
; TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/848,252
; FILING DATE: 29-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/06/538,052
; FILING DATE:

; APPLICATION NUMBER: US 08/151,672
; FILING DATE: 15-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Mcdiarmid, Shona S.
; REGISTRATION NUMBER: P-38,798
; REFERENCE/DOCKET NUMBER: 3158-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; TELEX: 06-23115
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-848-252-4

Query Match 0.7%; Score 8; DB 1; Length 76;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLLALL 1159
|||||||
DB 12 GLLLALL 19

RESULT 22
US-09-083-521-5
; Sequence 5, Application US/09083521
; Patent No. 6048970
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: PROSTATE GROWTH-ASSOCIATED MEMBRANE PROTEINS
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/083,521
; FILING DATE: Herewith
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: CERRONE, MICHAEL C.
; REGISTRATION NUMBER: 39,132
; REFERENCE/DOCKET NUMBER: PF-0527 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 76 amino acids
; TYPE: amino acid
; STRANDEDNESS: Single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1216498
US-09-083-521-5

Query Match 0.7%; Score 8; DB 3; Length 76;
Best Local Similarity 100.0%; Pred. No. 11;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALL 1159
Db 12 GLLALL 19

RESULT 23
US-08-848-252-2
; Sequence 2, Application US/08848252
; Patent No. 5804177
; GENERAL INFORMATION:
; APPLICANT: Humphries, Keith R.
; TITLE OF INVENTION: METHOD OF USING CD24 AS A CELL MARKER
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bereskin & Parr
; STREET: 40 King Street West
; CITY: Toronto
; STATE: Ontario
; COUNTRY: Canada
; ZIP: M5H 3Y2
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/848,252
; FILING DATE: 29-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/538,052
; FILING DATE:
; APPLICATION NUMBER: US 08/151,672
; FILING DATE: 15-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: McDiarmid, Shona S.
; REGISTRATION NUMBER: P-38,798
; REFERENCE/DOCKET NUMBER: 3158-028
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (416) 364-7311
; TELEFAX: (416) 361-1398
; TELEX: 06-23115
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 80 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-848-252-2

Query Match 0.7%; Score 8; DB 1; Length 80;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1152 GLLALL 1159
Db 12 GLLALL 19

RESULT 24
US-07-626-618A-2
; Sequence 2, Application US/07626618A
; Patent No. 5422265
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5422265el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/626,618A
FILING DATE: 7 DEC 1990
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5422265nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 810-221-8317
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-07-626-618A-2

Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 25
US-07-928-611-2
; Sequence 2, Application US/07928611
; Patent No. 5569601
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/928,611
FILING DATE: 19920810
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: No. 5569601nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-B
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-715-1000
TELEFAX: 312-715-1234
TELEX: 810-221-8317
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:

;
;
; LENGTH: 95 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-928-611-2

Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
Db 81 LLLALLVL 88

RESULT 26
US-08-333-977-2
; Sequence 2, Application US/08333977
; Patent No. 5594108
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,977
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/626,618
; FILING DATE: 7 DEC 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5594108nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 2:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-333-977-2

Query Match 0.7%; Score 8; DB 1; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
Db 81 LLLALLVL 88

RESULT 27
US-08-487-811A-2
; Sequence 2, Application US/08487811A
; Patent No. 5883226
; GENERAL INFORMATION:

;
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,811A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5883226nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:

INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 95 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-811A-2

Query Match 0.7%; Score 8; DB 2; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
| | | | |
Db 81 LLLALLVL 88

RESULT 28
US-09-060-694-2
; Sequence 2, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606

COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303

```
/ REFERENCE/DOCKET NUMBER: 90,1092-MM
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-913-0001
/ TELEFAX: 312-913-0002
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 95 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-09-060-694-2

Query Match      0.7%; Score 8; DB 3; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 29
US-09-378-074-2
/ Sequence 2, Application US/09378074
/ Patent No. 6437114
/ GENERAL INFORMATION:
/ APPLICANT: Van Tol, Hubert H.M.
/ TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Allegretti & Witcoff, Ltd.
/ STREET: 10 South Wacker Drive, Suite 3000
/ CITY: Chicago
/ STATE: Illinois
/ COUNTRY: USA
/ ZIP: 60606
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/378,074
/ FILING DATE: 20-Aug-1999
/ CLASSIFICATION: <Unknown>
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 07/928,611
/ FILING DATE: <Unknown>
/ ATTORNEY/AGENT INFORMATION:
/ NAME: No. 6437114han, Kevin E
/ REGISTRATION NUMBER: 35,303
/ REFERENCE/DOCKET NUMBER: 90,1092-B
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-715-1000
/ TELEFAX: 312-715-1234
/ TELEX: 810-221-8317
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 95 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-378-074-2

Query Match      0.7%; Score 8; DB 4; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

REFERENCE/DOCKET NUMBER: 90,1092-MM
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 312-913-0001
/ TELEFAX: 312-913-0002
/ TELEX:
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 95 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
US-09-060-694-2

Query Match      0.7%; Score 8; DB 3; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 30
PCT-US93-07370-2
/ Sequence 2, Application PC/TUS9307370
/ GENERAL INFORMATION:
/ APPLICANT:
/ TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
/ NUMBER OF SEQUENCES: 22
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: PCT/US93/07370
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 95 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
PCT-US93-07370-2

Query Match      0.7%; Score 8; DB 5; Length 95;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 31
US-08-973-544-4
/ Sequence 4, Application US/08973544
/ Patent No. 6338950
/ GENERAL INFORMATION:
/ APPLICANT: WEISS, Elisabeth
/ TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1
/ NUMBER OF SEQUENCES: 4
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NIKAIKO, MARCELSTEIN, MURRAY & ORAM LLP
/ STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby
/ CITY: Washington
/ STATE: DC
/ COUNTRY: USA
/ ZIP: 20005-5701
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/973,544
/ FILING DATE: 18-DEC-1997
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: PCT EP 96/02663
/ FILING DATE: 20-JUN-1996
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: EP 95109511.6
/ FILING DATE: 20-JUN-1995
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: EP 95112201.9
/ FILING DATE: 03-AUG-1995
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kitts, Monica Chin
/ REGISTRATION NUMBER: 36,105
/ REFERENCE/DOCKET NUMBER: P8341-7073
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (202) 638-5000
```

TELEFAX: (202) 638-4810
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 97 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-973-544-4

Query Match 0.7%; Score 8; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLLA 1157
Db 16 LGGLLLLA 23

RESULT 32
US-08-973-544-3
; Sequence 3, Application US/08973544
; Patent No. 6338950
; GENERAL INFORMATION:
; APPLICANT: WEISS, Elisabeth
; TITLE OF INVENTION: NEW IMMUNOREGULATORY PROTEIN LST-1
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIKALDO, MARMELESTEIN, MURRAY & ORAM LLP
; STREET: 655 Fifteenth St., NW, Suite 300, G St. Lobby
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/973,544
; FILING DATE: 18-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT EP 96/02663
; FILING DATE: 20-JUN-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95109511.6
; FILING DATE: 20-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95112201.9
; FILING DATE: 03-AUG-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kitts, Monica Chin
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: P8341-7073
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-5000
; TELEFAX: (202) 638-4810
; INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 104 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-973-544-3

Query Match 0.7%; Score 8; DB 4; Length 104;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGLLLLA 1157
Db 16 LGGLLLLA 23

Db 16 LGGLLLLA 23

RESULT 33
US-09-489-039A-10236
; Sequence 10236, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10236
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10236

Query Match 0.7%; Score 8; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
Db 13 LLLALLV 20

RESULT 34
US-09-419-568F-28
; Sequence 28, Application US/09419568F
; Patent No. 6331613
; GENERAL INFORMATION:
; APPLICANT: Dumoutier, Laure
; APPLICANT: Louhed, Jamila
; APPLICANT: Renaud, Jean-Christophe
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fac
; TITLE OF INVENTION: (TIFs) The Proteins Encoded, and Uses Thereof
; FILE REFERENCE: LUD 5543.2
; CURRENT APPLICATION NUMBER: US/09/419,568F
; CURRENT FILING DATE: 1999-10-18
; PRIOR APPLICATION NUMBER: US09/354,243
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US09/178,973
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 28
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-419-568F-28

Query Match 0.7%; Score 8; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
Db 21 LLLALLV 28

RESULT 35
US-09-354-243B-28
; Sequence 28, Application US/09354243B
; Patent No. 6359117
; GENERAL INFORMATION:
; APPLICANT: Dumoutier, Laure
; APPLICANT: Louhed, Jamila

```

; APPLICANT: Renault, Jean-Christophe
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules which Encode T Cell Inducible Fa
; TITLE OF INVENTION: (RIFs)
; TITLE OF INVENTION: The Proteins Encoded, and Uses Thereof
; FILE REFERENCE: LUD 5543.1
; CURRENT APPLICATION NUMBER: US/09/354,243B
; CURRENT FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: US09/178,973
; PRIOR FILING DATE: 1998-10-26
; NUMBER OF SEQ ID NOS: 29
; SEQ ID NO 28
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
US-09-354-243B-28

Query Match      0.7%; Score 8; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
   |||||
Db 21 LLLLLLV 28

RESULT 36
US-09-870-574-2
; Sequence 2, Application US/09870574
; Patent No. 6551799
; GENERAL INFORMATION:
; APPLICANT: Gurney, Austin L.
; APPLICANT: Aggarwal, Sudeepa
; APPLICANT: Xie, Ming-Hong
; APPLICANT: Maruoka, Ellen M.
; APPLICANT: Foster, Jessica S.
; APPLICANT: Goddard, Audrey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: INTERLEUKIN-22 POLYPEPTIDES, NUCLEIC ACIDS ENCODING
; TITLE OF INVENTION: THE SAME AND METHODS FOR THE TREATMENT OF PANCREATIC DISORDERS
; FILE REFERENCE: P2806-1 (US)
; CURRENT APPLICATION NUMBER: US/09/870,574
; CURRENT FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: US 60/169,495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/23328
; PRIOR FILING DATE: 2000-08-24
; NUMBER OF SEQ ID NOS: 7
; SEQ ID NO 2
; LENGTH: 179
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-870-574-2

Query Match      0.7%; Score 8; DB 4; Length 179;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLLLLV 1160
   |||||
Db 21 LLLLLLV 28

RESULT 37
US-09-252-991A-25905
; Sequence 25905, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
```

```

; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25905
; LENGTH: 183
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25905

Query Match      0.7%; Score 8; DB 4; Length 183;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1171 SARRRREP 1178
   |||||
Db 21 SARRRREP 28

RESULT 38
US-09-040-483-5
; Sequence 5, Application US/09040483
; Patent No. 6143867
; GENERAL INFORMATION:
; APPLICANT: Aketdiom, Ingrid E.
; TITLE OF INVENTION: NOVEL HUMAN EOSINOPHIL-
; TITLE OF INVENTION: DERIVED BASIC PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Icyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/040,483
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/740,036
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0133 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 544241
US-09-040-483-5

Query Match      0.7%; Score 8; DB 3; Length 234;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1153 LLLALLV 1160
Db 4 LLLALLV 11

RESULT 39
US-08-740-036-5
; Sequence 5, Application US/08740036
; Patent No. 6452507
; GENERAL INFORMATION:
; APPLICANT: Aketblom, Ingrid E.
; TITLE OF INVENTION: NOVEL HUMAN EOSINOPHIL-
; DERIVED BASIC PROTEIN
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/740,036
; FILING DATE: Filed Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0133 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 234 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 544241
; US-08-740-036-5

Query Match 0.7%; Score 8; DB 4; Length 234;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1153 LLLALLV 1160
Db 4 LLLALLV 11

RESULT 40
US-08-118-270-28
; Sequence 28, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-118-270-28

Query Match 0.7%; Score 8; DB 1; Length 315;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLV 1161
Db 48 LLLALLV 55

RESULT 41
PCT-US93-08528-28
; Sequence 28, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY=2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197


```
;
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 315 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
PCT-US93-08528-28

Query Match 0.7%; Score 8; DB 5; Length 315;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 48 LLLALLVL 55

RESULT 42
US-09-489-039A-7417
; Sequence 7417, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 7417
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-7417

Query Match 0.7%; Score 8; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1151 GGLLLAL 1158
Db 76 GGLLLAL 83

RESULT 43
US-09-046-736-4
; Sequence 4, Application US/09046736
; Patent No. 6090582
; GENERAL INFORMATION:
; APPLICANT: KIKUY, KRISTINE
; APPLICANT: ERICKSON-MILLER, CONNIE
; TITLE OF INVENTION: Sialoadhesin Family Member-3
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Ratner & Prestia
; STREET: P.O. Box 980
; CITY: Valley Forge
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046,736
; FILING DATE: 24-MAR-1998
```

```
;
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/041,885
; FILING DATE: 02-APR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Prestia, Paul F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-50019
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 374 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-046-736-4

Query Match 0.7%; Score 8; DB 3; Length 374;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 GVLGAVG 386
Db 257 GVLGAVG 264

RESULT 44
US-07-626-618A-17
; Sequence 17, Application US/07626618A
; Patent No. 5422265
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5422265el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/626,618A
; FILING DATE: 7 DEC 1990
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5422265nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-07-626-618A-17

Query Match 0.7%; Score 8; DB 1; Length 387;
```

```
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 45
US-08-086-439C-3
; Sequence 3, Application US/08086439C
; Patent No. 5468615
; GENERAL INFORMATION:
; APPLICANT: Chio, Christopher L.
; APPLICANT: Huff, Rita M.
; TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Upjohn Company, Corp. Intellectual
; ADDRESSEE: Property Law
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/086,439C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Darnley Jr., James D.
; REGISTRATION NUMBER: 33,673
; REFERENCE/DOCKET NUMBER: 4700
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 616-385-5210
; TELEFAX: 616-385-6897
; TELEX: 224401
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-086-439C-3

Query Match 0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 46
US-08-056-051-2
; Sequence 2, Application US/08056051
; Patent No. 5516683
; GENERAL INFORMATION:
; APPLICANT: Grandy, David K
; APPLICANT: Bunzow, James R
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.-M.
; TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,611
; FILING DATE: 19920810
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5569601nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
```

```
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/056,051
; FILING DATE: 19930429
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5516683nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 910-221-5317
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-056-051-2

Query Match 0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 47
US-07-928-611-18
; Sequence 18, Application US/07928611
; Patent No. 5569601
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.-M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/928,611
; FILING DATE: 19920810
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5569601nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
```

```
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 387 amino acids
;   TYPE: AMINO ACID
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-928-611-18

Query Match      0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 48
US-08-333-977-17
; Sequence 17, Application US/08333977
; Patent No. 5594108
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 5594108el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/333,977
; FILING DATE: 03-NOV-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/626,618
; FILING DATE: 7 DEC 1990
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5594108nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 387 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-333-977-17

Query Match      0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 49
US-08-449-153-1
; Sequence 1, Application US/08449153
; Patent No. 5686255
; GENERAL INFORMATION:
; APPLICANT: Deth, Richard C.
; TITLE OF INVENTION: Compositions and Methods for Diagnosing
; TITLE OF INVENTION: Schizophrenia
; NUMBER OF SEQUENCES: 1
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
; STREET: Two Militia Drive
; CITY: Lexington
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02173
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,153
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Carroll, Alice O.
; REGISTRATION NUMBER: 33,542
; REFERENCE/DOCKET NUMBER: RCD95-02
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-861-6240
; TELEFAX: 617-861-9540
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 387 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-449-153-1

Query Match      0.7%; Score 8; DB 1; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 50
US-08-434-877-3
; Sequence 3, Application US/08434877
; Patent No. 5721132
; GENERAL INFORMATION:
; APPLICANT: Chio, Christopher L.
; APPLICANT: Huff, Rita M.
; TITLE OF INVENTION: A Synthetic Gene for D4 Dopamine
; TITLE OF INVENTION: Receptors
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Upjohn Company, Corp. Intellectual
; ADDRESS: Property Law
; STREET: 301 Henrietta Street
; CITY: Kalamazoo
; STATE: Michigan
; COUNTRY: USA
; ZIP: 49001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette (DS,HD)
; COMPUTER: Gateway 2000, P5-90
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,877
```

; FILING DATE: 1 July 1993
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dainley Jr., James D.
 ; REGISTRATION NUMBER: 33,673
 ; REFERENCE/DOCKET NUMBER: 4700 DVI
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 616-385-5210
 ; TELEFAX: 616-385-6897
 ; TELEX: 224401
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 387 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-434-877-3

Query Match 0.7%; Score 8; DB 1; Length 387;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
 Db 81 LLLALLVL 88

RESULT 51
 US-08-449-651-1
 ; Sequence 1, Application US/08449651
 ; Patent No. 5738998
 ; GENERAL INFORMATION:
 ; APPLICANT: Deth, Richard C.
 ; TITLE OF INVENTION: Compositions and Methods for Diagnosing
 ; TITLE OF INVENTION: Schizophrenia
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 ; STREET: Two Militia Drive
 ; CITY: Lexington
 ; STATE: Massachusetts
 ; COUNTRY: USA
 ; ZIP: 02173
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/449,651
 ; FILING DATE: 24-MAY-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Carroll, Alice O.
 ; REGISTRATION NUMBER: 33,542
 ; REFERENCE/DOCKET NUMBER: RCD95-02
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617-861-6240
 ; TELEFAX: 617-861-9540
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 387 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-449-651-1

Query Match 0.7%; Score 8; DB 1; Length 387;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
 Db 81 LLLALLVL 88

RESULT 52
 US-08-487-811A-18
 ; Sequence 18, Application US/08487811A
 ; Patent No. 5883226
 ; GENERAL INFORMATION:
 ; APPLICANT: Civelli, Olivier
 ; APPLICANT: Van Tol, Robert H.M.
 ; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
 ; NUMBER OF SEQUENCES: 24
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
 ; STREET: 300 South Wacker Drive
 ; CITY: Chicago
 ; STATE: IL
 ; COUNTRY: USA
 ; ZIP: 60606
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/487,811A
 ; FILING DATE: 07-JUN-1995
 ; CLASSIFICATION: 530
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: No. 5883226nat, Kevin E
 ; REGISTRATION NUMBER: 35,303
 ; REFERENCE/DOCKET NUMBER: 90,1092-L
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312-913-0001
 ; TELEFAX: 312-913-0002
 ; TELEX:

; INFORMATION FOR SEQ ID NO: 18:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 387 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-487-811A-18
 Query Match 0.7%; Score 8; DB 2; Length 387;
 Best Local Similarity 100.0%; Pred. No. 48;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
 Db 81 LLLALLVL 88

RESULT 53
 US-08-833-703-1
 ; Sequence 1, Application US/08833703
 ; Patent No. 6080549
 ; GENERAL INFORMATION:
 ; APPLICANT: DETH, RICHARD
 ; TITLE OF INVENTION: METHODS AND MATERIALS FOR THE DIAGNOSIS AND
 ; TITLE OF INVENTION: TREATMENT OF SCHIZOPHRENIA AND RELATED DISORDERS
 ; NUMBER OF SEQUENCES: 1
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes
 ; STREET: Ten Post Office Square
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02109
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette

```
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/833,703
; FILING DATE: 08-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Heine, Ph.D., Holliday C
; REGISTRATION NUMBER: 34,346
; REFERENCE/DOCKET NUMBER: NU-431XX
; TELEPHONE: 617-542 2290
; TELEFAX: 617-451 0313
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-833-703-1

Query Match 0.7%; Score 8; DB 3; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 54
US-08-475-742-4
; Sequence 4, Application US/08475742
; Patent No. 6121015
; GENERAL INFORMATION:
; APPLICANT: O'Malley, Karen L
; APPLICANT: Todd, Richard D
; TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4 Receptor
; FILE REFERENCE: WU 102 CON DIV
; CURRENT APPLICATION NUMBER: US/08/475,742
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: US 08/261,293
; EARLIER FILING DATE: 1994-06-16
; EARLIER APPLICATION NUMBER: US 08/014,013
; EARLIER FILING DATE: 1993-01-28
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 4
; LENGTH: 387
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: PEPTIDE
; LOCATION: (1)..(387)
; OTHER INFORMATION: Human D4 Receptor Protein
; US-08-475-742-4

Query Match 0.7%; Score 8; DB 3; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 55
US-09-060-694-18
; Sequence 18, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/060,694
; FILING DATE: 15-APR-1998
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6203998nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-MM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-060-694-18

Query Match 0.7%; Score 8; DB 3; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 56
US-09-378-074-18
; Sequence 18, Application US/09378074
; Patent No. 6437114
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civelli, Olivier
; TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/09/378,074
; FILING DATE: 20-Aug-1999
; CLASSIFICATION: <Unknown>
```

```

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,611
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 643711nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 18:
US-09-378-074-18

Query Match 0.7%; Score 8; DB 4; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 57
US-08-261-293-4
; Sequence 4, Application US/08261293
; Patent No. 6486310
; GENERAL INFORMATION:
; APPLICANT: O'Malley, Karen L.
; APPLICANT: Todd, Richard D.
; TITLE OF INVENTION: Gene Encoding the Rat Dopamine D4
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/261,293
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/014,013
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Patst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: 31,284
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6524
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO

```

```

; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: Protein
; LOCATION: 1..387
; OTHER INFORMATION: /note= "Human D4 Receptor Protein"
US-08-261-293-4

Query Match 0.7%; Score 8; DB 4; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 58
PCT-US93-07370-18
; Sequence 18, Application PC/TUS9307370
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07370
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US93-07370-18

Query Match 0.7%; Score 8; DB 5; Length 387;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 59
US-08-056-051-4
; Sequence 4, Application US/08056051
; Patent No. 5516683
; GENERAL INFORMATION:
; APPLICANT: Grandy, David K
; APPLICANT: Bunzow, James R
; APPLICANT: Civeilli, Olivier
; APPLICANT: Van Tol, Hubert H.-M.
; TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/056,051
;; FILING DATE: 19930429
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 5516683nan, Kevin E
;; REGISTRATION NUMBER: 35,303
;; REFERENCE/DOCKET NUMBER: 90,1092-C
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312-715-1000
;; TELEFAX: 312-715-1234
;; TELEX: 910-221-5317
;; INFORMATION FOR SEQ ID NO: 4:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 419 amino acids
;; TYPE: AMINO ACID
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-08-056-051-4

Query Match 0.7%; Score 8; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 60
US-07-928-611-20
; Sequence 20, Application US/07928611
; Patent No. 5569601
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; APPLICANT: Civeilli, Olivier
; TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 19920810
; APPLICATION NUMBER: US/07/928,611
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5569601nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-928-611-20

Query Match 0.7%; Score 8; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 61
US-08-487-811A-20
; Sequence 20, Application US/08487811A
; Patent No. 5883226
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,811A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 5883226nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-L
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-913-0001
; TELEFAX: 312-913-0002
; TELEX:
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-487-811A-20

Query Match 0.7%; Score 8; DB 2; Length 419;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
Db 81 LLLALLVL 88

RESULT 62
US-09-060-694-20
; Sequence 20, Application US/09060694
; Patent No. 6203998
; GENERAL INFORMATION:
; APPLICANT: Civeilli, Olivier
; APPLICANT: Van Tol, Hubert H.M.
; TITLE OF INVENTION: A No. 6203998el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: McDonnell Boenhen Hulbert & Berghoff
; STREET: 300 South Wacker Drive
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:

;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.25
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/09/060,694
;; FILING DATE: 15-APR-1998
;; CLASSIFICATION: 530
;; ATTORNEY/AGENT INFORMATION:
;; NAME: No. 6203998nan, Kevin E
;; REGISTRATION NUMBER: 35,303
;; REFERENCE/DOCKET NUMBER: 90,1092-MM
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 312-913-0001
;; TELEFAX: 312-913-0002
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 20:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 419 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
US-09-060-694-20

Query Match 0.7%; Score 8; DB 3; Length 419;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
|||
DB 81 LLLALLVL 88

RESULT 63
US-09-378-074-20
; Sequence 20, Application US/09378074
; Patent No. 6437114
; GENERAL INFORMATION:
; APPLICANT: Van Tol, Hubert H.M.
; Civeilli, Olivier
; TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Allegretti & Witcoff, Ltd.
; STREET: 10 South Wacker Drive, Suite 3000
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60606
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/378,074
; FILING DATE: 20-Aug-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/928,611
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6437114nan, Kevin E
; REGISTRATION NUMBER: 35,303
; REFERENCE/DOCKET NUMBER: 90,1092-B
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-715-1000
; TELEFAX: 312-715-1234
; TELEX: 810-221-8317
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid

;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-378-074-20

Query Match 0.7%; Score 8; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
|||
DB 81 LLLALLVL 88

RESULT 64
PCT-US93-07370-20
; Sequence 20, Application PC/TUS9307370
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Novel Human Dopamine Receptor and Uses
; NUMBER OF SEQUENCES: 22
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07370
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 419 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US93-07370-20

Query Match 0.7%; Score 8; DB 5; Length 419;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
|||
DB 81 LLLALLVL 88

RESULT 65
US-09-252-991A-17394
; Sequence 17394, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 17394
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-17394

Query Match 0.7%; Score 8; DB 4; Length 453;
Best Local Similarity 100.0%; Pred. No. 56;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1150 LGGILLLA 1157
|||

Db 426 LGGLLLLA 433

RESULT 66

US-08-056-051-6

Sequence 6, Application US/08056051

Patent No. 5516683

GENERAL INFORMATION:

APPLICANT: Grandy, David K

APPLICANT: Bunzow, James R

APPLICANT: Civeilli, Olivier

APPLICANT: Van Tol, Hubert H.M.

TITLE OF INVENTION: A No. 5516683el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/056,051

FILING DATE: 19930429

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: No. 5516683nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-C

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-715-1000

TELEFAX: 312-715-1234

TELEX: 910-221-5317

INFORMATION FOR SEQ ID NO: 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: AMINO ACID

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-056-051-6

Query Match 0.7%; Score 8; DB 1; Length 467;

Best Local Similarity 100.0%; Pred. No. 57;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

RESULT 67

US-07-928-611-22

Sequence 22, Application US/07928611

Patent No. 5569601

GENERAL INFORMATION:

APPLICANT: Van Tol, Hubert H.M.

APPLICANT: Civeilli, Olivier

TITLE OF INVENTION: A No. 5569601el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Allegretti & Witcoff, Ltd.

STREET: 10 South Wacker Drive, Suite 3000

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,811A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: No. 5883226nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-L

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-487-811A-22

Qy 1154 LLLALLVL 1161

Db 81 LLLALLVL 88

RESULT 68

US-08-487-811A-22

Sequence 22, Application US/08487811A

Patent No. 5883226

GENERAL INFORMATION:

APPLICANT: Civeilli, Olivier

APPLICANT: Van Tol, Hubert H.M.

TITLE OF INVENTION: A No. 5883226el Human Dopamine Receptor and Uses

NUMBER OF SEQUENCES: 24

CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff

STREET: 300 South Wacker Drive

CITY: Chicago

STATE: IL

COUNTRY: USA

ZIP: 60606

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/487,811A

FILING DATE: 07-JUN-1995

CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: No. 5883226nan, Kevin E

REGISTRATION NUMBER: 35,303

REFERENCE/DOCKET NUMBER: 90,1092-L

TELECOMMUNICATION INFORMATION:

TELEPHONE: 312-913-0001

TELEFAX: 312-913-0002

TELEX:

INFORMATION FOR SEQ ID NO: 22:

SEQUENCE CHARACTERISTICS:

LENGTH: 467 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-487-811A-22

Query Match 0.7%; Score 8; DB 2; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 69

US-09-046-736-2
Sequence 2, Application US/09046736
Patent No. 6090582
GENERAL INFORMATION:
APPLICANT: KIRBY, KRISTINE
TITLE OF INVENTION: Sialoadhesin Family Member-3
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia
STREET: P.O. Box 980
CITY: Valley Forge
STATE: PA
COUNTRY: USA
ZIP: 19482

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/046,736
FILING DATE: 24-MAR-1998
CLASSIFICATION: 514

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/041,885
FILING DATE: 02-APR-1997
ATTORNEY/AGENT INFORMATION:
NAME: Prestia, Paul F
REGISTRATION NUMBER: 23,031
REFERENCE/DOCKET NUMBER: GH-50019
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-407-0700
TELEFAX: 610-407-0701
TELEX:

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-046-736-2

Query Match 0.7%; Score 8; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 379 GVLGAVG 386
DB 350 GVLGAVG 357

RESULT 70

US-09-060-694-22
Sequence 2, Application US/09060694
Patent No. 6203998
GENERAL INFORMATION:
APPLICANT: Civeilli, Olivier
TITLE OF INVENTION: Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 24
CORRESPONDENCE ADDRESS:

ADDRESSEE: McDonnell Boehnen Hulbert & Berghoff
STREET: 300 South Wacker Drive
CITY: Chicago
STATE: IL
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/060,694
FILING DATE: 15-APR-1998

CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: No. 6203998nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-MM
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-913-0001
TELEFAX: 312-913-0002
TELEX:

INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 467 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-060-694-22

Query Match 0.7%; Score 8; DB 3; Length 467;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1154 LLLALLVL 1161
DB 81 LLLALLVL 88

RESULT 71

US-09-378-074-22
Sequence 22, Application US/09378074
Patent No. 6437114
GENERAL INFORMATION:
APPLICANT: Van Tol, Hubert H.M.
Civeilli, Olivier

TITLE OF INVENTION: A No. 6437114el Human Dopamine Receptor and Uses
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Allegretti & Witcoff, Ltd.
STREET: 10 South Wacker Drive, Suite 3000
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60606

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/378,074
FILING DATE: 20-Aug-1999

CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/928,611
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: No. 6437114nan, Kevin E
REGISTRATION NUMBER: 35,303
REFERENCE/DOCKET NUMBER: 90,1092-B
TELECOMMUNICATION INFORMATION:

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,618
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:
NAME: Williams Jk., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32797
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-618-103

Query Match      0.7%; Score 8; DB 1; Length 494;
Best Local Similarity 100.0%; Pred.No. 60;
Matches      8; Conservative      0; Mismatches      0; Indels      0; Gaps      0;

QY      1151 GGLLLALL 1158
DB      442 GGLLLALL 449

RESULT 74
US-08-605-672-103
Sequence 103 Application US/08605672
Patent No. 5817515
GENERAL INFORMATION:
APPLICANT: Gallatin, W. Michael
APPLICANT: Van der Vieren, Monica
TITLE OF INVENTION: No. 5817515el Human 2 Integrin Alpha Subunit
NUMBER OF SEQUENCES: 103
CORRESPONDENCE ADDRESS:
ADDRESS: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 233 South Wacker Drive, 6300 Sear Tower
CITY: Chicago
STATE: Illinois
COUNTRY: United States
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/605,672
FILING DATE:
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/173,497
FILING DATE: 23-DEC-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/286,889
FILING DATE: 5-AUG-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,652
FILING DATE: 21-DEC-1994
ATTORNEY/AGENT INFORMATION:

```

NAME: Williams Jr., Joseph A.
REGISTRATION NUMBER: 38,659
REFERENCE/DOCKET NUMBER: 27866/32684
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312-474-6300
TELEFAX: 312-474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 103:
SEQUENCE CHARACTERISTICS:
LENGTH: 494 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-605-672-103

Query Match 0.7%; Score 8; DB 2; Length 494;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1151 GGLLLAL 1158
| | | | |
Db 442 GGLLLAL 449

RESULT 75

US-08-482-293A-103
; Sequence 103, Application US/08482293A
; Patent No. 5831029
; GENERAL INFORMATION:
; APPLICANT: Gallatin, W. Michael
; APPLICANT: Van der Vieren, Monica
; TITLE OF INVENTION: No. 5831029el Human 2 Integrin Alpha Subunit
; NUMBER OF SEQUENCES: 103
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive, 6300 Sear Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,293A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/173,437
; FILING DATE: 23-DEC-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/286,889
; FILING DATE: 5-AUG-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,652
; FILING DATE: 21-DEC-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Williams Jr., Joseph A.
; REGISTRATION NUMBER: 38,659
; REFERENCE/DOCKET NUMBER: 27866/32684
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-474-6300
; TELEFAX: 312-474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 103:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 494 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-482-293A-103

Query Match 0.7%; Score 8; DB 2; Length 494;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1151 GGLLLAL 1158
| | | | |
Db 442 GGLLLAL 449

Search completed: September 21, 2004, 13:42:57
Job time : 24 secs